

The Length of the Longest Common Subsequence of Two Independent Mallows Permutations

Ke Jin*

January 30, 2017

Abstract

The Mallows measure is a probability measure on S_n where the probability of a permutation π is proportional to $q^{l(\pi)}$ with $q > 0$ being a parameter and $l(\pi)$ the number of inversions in π . We prove a weak law of large numbers for the length of the longest common subsequences of two independent permutations drawn from the Mallows measure, when q is a function of n and $n(1 - q)$ has limit in \mathbb{R} as $n \rightarrow \infty$.

1 Introduction

1.1 Background

The longest common subsequence(LCS) problem is a classical problem which has application in fields such as molecular biology (see, e.g., [20]) , data comparison and software version control. Most previous works on the LCS problem are focused on the case when the strings are generated uniformly at random from a given alphabet. Notably, Chvátal and Sankoff [4] proved that the expected length of the LCS of two random k -ary sequences of length n when normalized by n converges to a constant γ_k . Since then, various endeavors [8, 6, 7, 16] have been made to determine the value of γ_k . The

*Department of Mathematical Sciences, University of Delaware. EMAIL: kejin@udel.edu

exact values of γ_k are still unknown. The known lower and upper bounds [16] for γ_2 are

$$0.788071 < \gamma_2 < 0.826280.$$

In contrast to the LCS of two random strings, the LCS of two permutations is well connected to the longest increasing subsequence(LIS) problem (cf. Proposition 3.1 in [12]). This can be seen from the following two facts,

- For any $\pi \in S_n$, the length of the LCS of π and the identity in S_n is equal to the length of the LIS of π .
- For any $\pi, \tau \in S_n$, the length of the LCS of π and τ is equal to the length of the LCS of $\tau^{-1} \circ \pi$ and the identity in S_n .

From the above two properties, it is easily seen that, if π, τ are independent and either π or τ is uniformly distributed on S_n the length of the LCS of π and τ has the same distribution as the length of the LIS of a uniformly random permutation. The length of the LIS of a uniformly random permutation has been well studied with major contributions from Hammersley [11], Logan and Shepp [15], Vershik and Kerov [14] and culminating with the groundbreaking work of Baik, Deift and Johansson [2] who prove that, under proper scaling, the length of the LIS converges to the Tracy-Widom distribution. Therefore, the length of the LCS of two permutations is only of interest when both permutations are non-uniformly distributed. In this paper we study the length of the LCS of two independent permutations drawn from the Mallows measure.

Definition 1.1. *Given $\pi \in S_n$, the inversion set of π is defined by*

$$\text{Inv}(\pi) := \{(i, j) : 1 \leq i < j \leq n \text{ and } \pi(i) > \pi(j)\},$$

and the inversion number of π , denoted by $l(\pi)$, is defined to be the cardinality of $\text{Inv}(\pi)$.

The Mallows measure on S_n is introduced by Mallows in [17]. For $q > 0$, the (n, q) - Mallows measure on S_n is given by

$$\mu_{n,q}(\pi) := \frac{q^{l(\pi)}}{Z_{n,q}},$$

where $Z_{n,q}$ is the normalizing constant. In other words, under the Mallows measure with parameter $q > 0$, the probability of a permutation π is proportional to $q^{l(\pi)}$. Mallows measure has been used in modeling ranked and partially ranked data (see, e.g., [5], [10], [18]).

Definition 1.2. For any $\pi, \tau \in S_n$, define the length of the longest common subsequence of π and τ as follows,

$$LCS(\pi, \tau) := \max(m : \exists i_1 < \dots < i_m \text{ and } j_1 < \dots < j_m \\ \text{such that } \pi(i_k) = \tau(j_k) \text{ for all } k \in [m]).$$

Given the close connection of the LCS of two permutations and the LIS problem, to prove our results, we are able to make use of the techniques developed in [19, 3] in which weak laws of large numbers of the length of the LIS of permutation under Mallows measure have been proven for different regimes of q .

1.2 Results

Before state the main theorem, we introduce the following lemma proved in [13], which shows the convergence of the empirical measure of a collection of random points defined by two independent Mallows permutations.

Lemma 1.3. Suppose that $\{q_n\}_{n=1}^\infty$ and $\{q'_n\}_{n=1}^\infty$ are two sequences such that $\lim_{n \rightarrow \infty} n(1-q_n) = \beta$ and $\lim_{n \rightarrow \infty} n(1-q'_n) = \gamma$, with $\beta, \gamma \in \mathbb{R}$. Let \mathbb{P}_n denote the probability measure on $S_n \times S_n$ such that $\mathbb{P}_n((\pi, \tau)) = \mu_{n,q_n}(\pi) \cdot \mu_{n,q'_n}(\tau)$, i. e. \mathbb{P}_n is the product measure of μ_{n,q_n} and μ_{n,q'_n} . For any $R = (x_1, x_2] \times (y_1, y_2] \subset [0, 1] \times [0, 1]$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\left| \frac{1}{n} \sum_{i=1}^n \mathbb{1}_R \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n} \right) - \int_R \rho(x, y) dx dy \right| > \epsilon \right) = 0, \quad (1)$$

for any $\epsilon > 0$, with

$$\rho(x, y) := \int_0^1 u(x, t, \beta) \cdot u(t, y, \gamma) dt. \quad (2)$$

Here

$$u(x, y, \beta) := \frac{(\beta/2) \sinh(\beta/2)}{(e^{\beta/4} \cosh(\beta[x-y]/2) - e^{-\beta/4} \cosh(\beta[x+y-1]/2))^2}, \quad (3)$$

for $\beta \neq 0$, and $u(x, y, 0) := 1$.

The density $u(x, y, \beta)$ in (3), obtained by Starr in [21], defines the limiting distribution of the empirical measure induced by Mallows permutation when the parameters q_n satisfy that $\lim_{n \rightarrow \infty} n(1 - q_n) = \beta$.

The main result of this paper is a weak law of large numbers of the LCS of two permutations drawn independently from the Mallows measure.

There are three main ingredients in our proof of the weak law of large numbers of the length of the LCS of two independent Mallows distributed permutations. The first observation, which is proved in Corollary 2.4, is that the length of LCS of two permutations π and τ is equal to the length of the longest increasing points in the collection of points

$$\mathbf{z}(\pi^{-1}, \tau^{-1}) := \left\{ \left(\frac{\pi^{-1}(i)}{n}, \frac{\tau^{-1}(i)}{n} \right) \right\}_{i \in [n]}.$$

The second observation, deduced from Lemma 1.3, is that the number of points in $\mathbf{z}(\pi^{-1}, \tau^{-1})$ contained in any fixed rectangle, when scaled by the size of the permutation, converges in probability to a constant. The third observation, proved in Lemma 4.4, is that the length of the longest increasing points in $\mathbf{z}(\pi^{-1}, \tau^{-1})$ within a small box R is close to the size of the LIS in the uniform case, i.e., it is approximately $2\sqrt{|\mathbf{z}(\pi^{-1}, \tau^{-1}) \cap R|}$. With these results, we prove our main theorem following the method developed by Deuschel and Zeitouni in [9] for record lengths of i.i.d. points.

Theorem 1.4. *Let B_{\nearrow}^1 denote the set of nondecreasing, C_b^1 functions $\phi : [0, 1] \rightarrow [0, 1]$, with $\phi(0) = 0$ and $\phi(1) = 1$. Define function $J : B_{\nearrow}^1 \rightarrow \mathbb{R}$,*

$$J(\phi) := \int_0^1 \sqrt{\dot{\phi}(x)\rho(x, \phi(x))} dx, \quad \text{and} \quad \bar{J} := \sup_{\phi \in B_{\nearrow}^1} J(\phi).$$

Here $\rho(x, y)$ is the density defined in (2). Under the same conditions as in Lemma 1.3, for any $\epsilon > 0$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\left| \frac{LCS(\pi, \tau)}{\sqrt{n}} - 2\bar{J} \right| < \epsilon \right) = 1. \quad (4)$$

Finally, we derive the limiting constant in the special case when $\beta = \gamma$.

Corollary 1.5. *Suppose that $\{q_n\}_{n=1}^\infty$ and $\{q'_n\}_{n=1}^\infty$ are two sequences such that $\lim_{n \rightarrow \infty} n(1 - q_n) = \lim_{n \rightarrow \infty} n(1 - q'_n) = \beta$ with $\beta \neq 0$. Then, the constant \bar{J} in Theorem 1.4 is given by*

$$\bar{J} = \sqrt{\frac{\beta}{6 \sinh(\beta/2)}} \cdot \int_0^1 \sqrt{\cosh(\beta/2) + 2 \cosh(\beta[2x - 1]/2)} dx.$$

2 Reduction LCS problem to LIS problem

Definition 2.1. Given a set of points in \mathbb{R}^2 : $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$, where $z_i = (x_i, y_i) \in \mathbb{R}^2$, we say that $(z_{i_1}, z_{i_2}, \dots, z_{i_m})$ is an increasing subsequence if

$$x_{i_j} < x_{i_{j+1}}, \quad y_{i_j} < y_{i_{j+1}}, \quad j = 1, 2, \dots, m-1.$$

Here we do not require $i_j < i_{j+1}$. Let $LIS(\mathbf{z})$ denote the length of the longest increasing subsequence of \mathbf{z} .

Definition 2.2. Given $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$, $\mathbf{b} = (b_1, \dots, b_n) \in \mathbb{R}^n$, we say that $((a_{i_1}, b_{i_1}), (a_{i_2}, b_{i_2}), \dots, (a_{i_m}, b_{i_m}))$ is an increasing subsequence between \mathbf{a} and \mathbf{b} if

$$a_{i_j} < a_{i_{j+1}}, \quad b_{i_j} < b_{i_{j+1}}, \quad j = 1, 2, \dots, m-1$$

Here we do not require $i_j < i_{j+1}$. Let $LIS(\mathbf{a}, \mathbf{b})$ denote the length of the longest increasing subsequence between \mathbf{a} and \mathbf{b} . Let $LIS(\mathbf{a}) := LIS(id, \mathbf{a})$, $LDS(\mathbf{a}) := LIS(id^r, \mathbf{a})$. Here $id = (1, 2, \dots, n)$ denotes the identity in S_n and $id^r = (n, \dots, 2, 1)$ denotes the reversal of identity in S_n . Hence $LIS(\mathbf{a})$ is the length of the longest increasing subsequence of \mathbf{a} and $LDS(\mathbf{a})$ is the length of the longest decreasing subsequence of \mathbf{a} .

Note that Definition 2.2 allows us to define $LIS(\pi, \tau)$, the length of the longest increasing subsequence of two permutations, by regarding π and τ as vectors in \mathbb{Z}^n . We show that $LCS(\pi, \tau) = LIS(\pi^{-1}, \tau^{-1})$, which allows us to reduce the LCS problem to an LIS problem.

Lemma 2.3. Given $\pi, \tau \in S_n$, we have

$$LCS(\pi, \tau) = LCS(\sigma \circ \pi, \sigma \circ \tau), \quad LIS(\pi, \tau) = LIS(\pi \circ \sigma, \tau \circ \sigma),$$

for any $\sigma \in S_n$.

Proof. Suppose (a_1, a_2, \dots, a_m) is a common subsequence of π and τ , then $(\sigma(a_1), \sigma(a_2), \dots, \sigma(a_m))$ is a common subsequence of $\sigma \circ \pi$ and $\sigma \circ \tau$. Hence,

$$LCS(\pi, \tau) \leq LCS(\sigma \circ \pi, \sigma \circ \tau) \leq LCS(\sigma^{-1} \circ \sigma \circ \pi, \sigma^{-1} \circ \sigma \circ \tau) = LCS(\pi, \tau).$$

Similarly, suppose $((\pi(i_1), \tau(i_1)), (\pi(i_2), \tau(i_2)), \dots, (\pi(i_m), \tau(i_m)))$ is an increasing subsequence between π and τ , then $((\pi \circ \sigma(i'_1), \tau \circ \sigma(i'_1)),$

$(\pi \circ \sigma(i'_2), \tau \circ \sigma(i'_2)), \dots, (\pi \circ \sigma(i'_m), \tau \circ \sigma(i'_m)))$ is an increasing subsequence between $\pi \circ \sigma$ and $\tau \circ \sigma$, where $i'_k = \sigma^{-1}(i_k)$ for $k \in [m]$. Hence,

$$\text{LIS}(\pi, \tau) \leq \text{LIS}(\pi \circ \sigma, \tau \circ \sigma) \leq \text{LIS}(\pi \circ \sigma \circ \sigma^{-1}, \tau \circ \sigma \circ \sigma^{-1}) = \text{LIS}(\pi, \tau).$$

□

Corollary 2.4. *For any $\pi, \tau \in S_n$, $\text{LCS}(\pi, \tau) = \text{LIS}(\pi^{-1}, \tau^{-1})$.*

Proof. By the previous lemma, we have

$$\text{LCS}(\pi, \tau) = \text{LCS}(id, \pi^{-1} \circ \tau) = \text{LIS}(id, \pi^{-1} \circ \tau) = \text{LIS}(\tau^{-1}, \pi^{-1})$$

In the second equality, we use the following trivial fact,

$$\text{LCS}(id, \pi) = \text{LIS}(\pi) = \text{LIS}(id, \pi)$$

Here, id denotes the identity in S_n , i.e. $id = (1, 2, \dots, n)$. □

3 Weak Bruhat order

To prove Lemma 4.4, which says that the LIS of the points $\{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i \in [n]}$ that fall in a small box is close to the uniform case, we will establish a coupling of permutations (X, Y, X', X'') such that given $\mathbf{a} = (a_1, \dots, a_k)$ with $a_i \in [n]$, $\text{LIS}(X_{\mathbf{a}}, Y_{\mathbf{a}})$ can be bounded by $\text{LIS}(X'_{\mathbf{a}})$ and $\text{LDS}(X''_{\mathbf{a}})$. Here X, X' and X'' are distributed according to $\mu_{n,q}$ and Y is independent of X with an arbitrary distribution on S_n . The main tool we use to construct the coupling is the Weak Bruhat order on S_n .

Recall that for a permutation $\pi \in S_n$, $l(\pi)$ denotes the number of inversions of π and $\text{Inv}(\pi)$ denotes the set of inversions of π .

Definition 3.1. *The left weak Bruhat order (S_n, \leq_L) is defined as the transitive closure of the relations*

$$\pi \leq_L \tau \quad \text{if} \quad \tau = s_i \circ \pi \quad \text{and} \quad l(\tau) = l(\pi) + 1,$$

where $s_i = (i, i+1)$ is an adjacent transposition in S_n .

We are multiplying permutations right-to-left. For instance, $s_2 \circ 2413 = 3412$. One characterization of the left weak order is the following (cf. [1]),

$$\pi \leq_L \tau \quad \text{if and only if} \quad \text{Inv}(\pi) \subseteq \text{Inv}(\tau).$$

The *right weak Bruhat order* (S_n, \leq_R) is defined in the same way except that the covering relationship is given by $\tau = \pi \circ s_i$ and $l(\tau) = l(\pi) + 1$.

Definition 3.2. Let (Ω, \leq) be a partially ordered set. A non-empty subset $A \subset \Omega$ is called increasing if

$$\omega \in A \text{ and } \omega \leq \omega' \Rightarrow \omega' \in A.$$

Given two probability measures μ_1, μ_2 on (Ω, \mathcal{F}) , we say that μ_1 is stochastically smaller than μ_2 , denoted by $\mu_1 \preceq \mu_2$, if

$$\mu_1(A) \leq \mu_2(A) \text{ for all increasing events } A.$$

Lemma 3.3. Given the poset (S_n, \leq_L) , for any $0 < q < q'$, we have $\mu_{n,q} \preceq \mu_{n,q'}$.

Proof. We are going to construct a coupling of two Markov chains (X_t, Y_t) , such that

1. Both $\{X_t\}$ and $\{Y_t\}$ are irreducible, aperiodic Markov chains on S_n .
2. The stationary distributions for $\{X_t\}$ and $\{Y_t\}$ are $\mu_{n,q}$ and $\mu_{n,q'}$ respectively.
3. $X_t \leq_L Y_t$ for any $t \geq 0$.

By 3, for any increasing subset $A \subseteq S_n$, we have

$$\mathbb{P}(X_t \in A) = \mathbb{P}(X_t \in A, X_t \leq_L Y_t) \leq \mathbb{P}(Y_t \in A). \quad (5)$$

Also, by properties 1 and 2, we have

$$\mu_{n,q}(A) = \lim_{t \rightarrow \infty} \mathbb{P}(X_t \in A), \quad \text{and} \quad \mu_{n,q'}(A) = \lim_{t \rightarrow \infty} \mathbb{P}(Y_t \in A). \quad (6)$$

Combining (5) and (6), we get

$$\mu_{n,q}(A) \leq \mu_{n,q'}(A), \quad \text{for all increasing subsets } A.$$

The remainder of the proof is devoted to the construction of the coupling (X_t, Y_t) which satisfies the three properties above. The coupling (X_t, Y_t) is defined as follows,

- When time $t = 0$, $X_0 = Y_0 = id$. That is, both chains start from the identity in S_n .

- At each time t , sample three independent random variables: U, F, B . U samples the integers from 1 to $n - 1$ uniformly. F and B are two coins with the probability of heads being $\frac{1}{1+q'}$ and $\frac{(1+q')q}{(1+q)q'}$ respectively. Suppose $U = i$. Then flip the coins F and B and update the chains according to the following rules:

Case 1 if $X^{-1}(i) < X^{-1}(i + 1)$ and $Y^{-1}(i) < Y^{-1}(i + 1)$, then

F is head	$X_{t+1} = X_t,$	$Y_{t+1} = Y_t$
F is tail, B is head	$X_{t+1} = s_i \circ X_t,$	$Y_{t+1} = s_i \circ Y_t$
F is tail, B is tail	$X_{t+1} = X_t,$	$Y_{t+1} = s_i \circ Y_t$

Case 2 if $X^{-1}(i) < X^{-1}(i + 1)$ and $Y^{-1}(i) > Y^{-1}(i + 1)$, then

F is head	$X_{t+1} = X_t,$	$Y_{t+1} = s_i \circ Y_t$
F is tail, B is head	$X_{t+1} = s_i \circ X_t,$	$Y_{t+1} = Y_t$
F is tail, B is tail	$X_{t+1} = X_t,$	$Y_{t+1} = Y_t$

Case 3 if $X^{-1}(i) > X^{-1}(i + 1)$ and $Y^{-1}(i) > Y^{-1}(i + 1)$, then

F is head	$X_{t+1} = s_i \circ X_t,$	$Y_{t+1} = s_i \circ Y_t$
F is tail, B is head	$X_{t+1} = X_t,$	$Y_{t+1} = Y_t$
F is tail, B is tail	$X_{t+1} = s_i \circ X_t,$	$Y_{t+1} = Y_t$

By the definition above and the following facts, it is easy to check that the three properties listed at the beginning of the proof are satisfied.

- The adjacent transpositions $\{s_i\}$ generate S_n under the group multiplication in S_n .
- If $\pi^{-1}(i) = j$, $\pi^{-1}(i + 1) = k$ and $j < k$, we have $\text{Inv}(s_i \circ \pi) = \text{Inv}(\pi) \cup \{(j, k)\}$.
- If $\pi^{-1}(i) = j$, $\pi^{-1}(i + 1) = k$ and $j > k$, we have $\text{Inv}(s_i \circ \pi) = \text{Inv}(\pi) \setminus \{(k, j)\}$.

- $\pi \leq_L \tau$ if and only if $\text{Inv}(\pi) \subseteq \text{Inv}(\tau)$.
- Both chains X_t, Y_t satisfy the detailed balance equations, i.e. that the Mallows distribution satisfies that

$$\mu_{n,q}(\pi) \cdot \mathbb{P}(X_{t+1} = \tau | X_t = \pi) = \mu_{n,q}(\tau) \cdot \mathbb{P}(X_{t+1} = \pi | X_t = \tau)$$

for any $\pi, \tau \in S_n$ and similarly for the chain Y_t .

□

Definition 3.4. Given $\pi \in S_n$ and $\mathbf{a} = (a_1, a_2, \dots, a_k)$, where $a_i \in [n]$ and $a_1 < a_2 < \dots < a_k$, let $\pi(\mathbf{a}) = (\pi(a_1), \pi(a_2), \dots, \pi(a_k))$. Let $\pi_{\mathbf{a}} \in S_k$ denote the permutation induced by $\pi(\mathbf{a})$, i. e. $\pi_{\mathbf{a}}(i) = j$ if $\pi(a_i)$ is the j -th smallest term in $\pi(\mathbf{a})$.

Corollary 3.5. Let $n \in \mathbb{N}$, and $\mathbf{a} = (a_1, a_2, \dots, a_k)$, where $a_i \in [n]$ and $a_1 < a_2 < \dots < a_k$.

- (a) For any $q \geq 1$, we can construct a pair of random variables (U, V) such that U is uniformly distributed on S_k , V has the same distribution as $\pi_{\mathbf{a}}$, where $\pi \sim \mu_{n,q}$, and $U \leq_L V$.
- (b) For any $q \leq 1$, we can construct a pair of random variables (U, V) such that U is uniformly distributed on S_k , V has the same distribution as $\pi_{\mathbf{a}}$, where $\pi \sim \mu_{n,q}$, and $V \leq_L U$.

Proof. Here we only prove part (a). Part (b) follows by a similar argument. Since $q \geq 1$, by Lemma 3.3 and Strassen's theorem [22], there exist two random variables (X, Y) defined on the same probability space such that X is the uniform measure on S_n , $Y \sim \mu_{n,q}$ and $X \leq_L Y$. Then we can show $X_{\mathbf{a}} \leq_L Y_{\mathbf{a}}$. Since $\pi \leq_L \tau$ if and only if $\text{Inv}(\pi) \subseteq \text{Inv}(\tau)$, we have

$$\begin{aligned} \text{Inv}(X_{\mathbf{a}}) &= \{(i, j) : 1 \leq i < j \leq k \text{ and } X(a_i) > X(a_j)\} \\ &\subset \{(i, j) : 1 \leq i < j \leq k \text{ and } Y(a_i) > Y(a_j)\} = \text{Inv}(Y_{\mathbf{a}}). \end{aligned}$$

Hence, if we define $U := X_{\mathbf{a}}$ and $V := Y_{\mathbf{a}}$, part (a) follows by the fact that $X_{\mathbf{a}}$ is uniformly distributed on S_k . □

Lemma 3.6. *Given $\pi, \tau \in S_k$ with $\pi \leq_L \tau$, for any $n \geq k$, $0 < q \leq 1$ and $a_1 < \dots < a_k$ with $a_i \in [n]$, there exists a coupling (X, Y) such that $X \sim \mu_{n,q}$, $Y \sim \mu_{n,q}$ and*

$$LIS(X_{\mathbf{a}}, \pi) \geq LIS(Y_{\mathbf{a}}, \tau).$$

Here $\mathbf{a} = (a_1, a_2, \dots, a_k)$.

Proof. It suffices to show the case when τ covers π in (S_k, \leq_L) , that is $l(\tau) = l(\pi) + 1$ and $\tau = s_i \circ \pi$ for some $i \in [k-1]$. Because suppose we have $\pi \leq_L \tau \leq_L \sigma$ in S_k and two couplings (X, Y) and (Y', Z) , which are not necessarily defined in the same probability space, such that X, Y, Y', Z have the same marginal distribution $\mu_{n,q}$ and

$$LIS(X_{\mathbf{a}}, \pi) \geq LIS(Y_{\mathbf{a}}, \tau), \quad LIS(Y'_{\mathbf{a}}, \tau) \geq LIS(Z_{\mathbf{a}}, \sigma). \quad (7)$$

We can construct a new coupling (X', Z') as follows,

- (1) Sample a permutation $\xi \in S_n$ according to the distribution $\mu_{n,q}$.
- (2) Sample X' according to the induced distribution on S_n by the first coupling (X, Y) conditioned on $Y = \xi$.
- (3) Sample Z' according to the induced distribution on S_n by the second coupling (Y', Z) conditioned on $Y' = \xi$.

By the definition of conditional probability, it is easily seen that $X' \sim \mu_{n,q}$ and $Z' \sim \mu_{n,q}$. Also, regardless of which permutation ξ being sampled in the first step, by (7), we have

$$LIS(X'_{\mathbf{a}}, \pi) \geq LIS(\xi_{\mathbf{a}}, \tau) \geq LIS(Z'_{\mathbf{a}}, \sigma).$$

In the remaining of the proof, we assume $\tau = s_i \circ \pi$ and $l(\tau) = l(\pi) + 1$. First we point out the following facts, for any $\sigma \in S_n$,

$$\sigma \circ (i, j) = (\sigma(i), \sigma(j)) \circ \sigma, \quad \sigma_{\mathbf{a}} \circ (i, j) = (\sigma \circ (a_i, a_j))_{\mathbf{a}}. \quad (8)$$

Here (i, j) denotes the transposition of i and j .

Let $r = a_{\pi^{-1}(i)}$ and $t = a_{\pi^{-1}(i+1)}$. Since $l(\tau) = l(\pi) + 1$, we have $\pi^{-1}(i) < \pi^{-1}(i+1)$, thus, $r < t$. Let $A := \{\{\sigma, \sigma \circ (r, t)\} : \sigma \in S_n \text{ and } \sigma(r) < \sigma(t)\}$. Clearly, A is a partition of S_n . Then we construct the coupling (X, Y) as follows:

- (1) Choose a set in A according to measure $\mu_{n,q}$, i.e. the set $\{\sigma, \sigma \circ (r, t)\}$ is chosen with probability $\mu_{n,q}(\{\sigma, \sigma \circ (r, t)\})$.
- (2) Suppose the set $\{\sigma, \sigma \circ (r, t)\}$, with $\sigma(r) < \sigma(t)$, is chosen in the first step. Flip a coin with probability of head equal to

$$p = \frac{q^{l(\sigma)} - q^{l(\sigma \circ (r, t))}}{q^{l(\sigma)} + q^{l(\sigma \circ (r, t))}}.$$

- (3) If the outcome is head, then we set $X = Y = \sigma$.
- (4) If the outcome is tail, then, with equal probability, we set either $X = \sigma$, $Y = \sigma \circ (r, t)$ or $X = \sigma \circ (r, t)$, $Y = \sigma$.

Here, in second step, the probability of head p is nonnegative. Because we have $0 < q \leq 1$ and the following fact:

$$i < j \text{ and } \sigma(i) < \sigma(j) \Rightarrow l(\sigma) < l(\sigma \circ (i, j)), \quad \forall \sigma \in S_n$$

It can be easily verified that (X, Y) thus defined have the correct marginal distribution $\mu_{n,q}$. In the following we show that

$$\text{LIS}(X_{\mathbf{a}} \circ \pi^{-1}) \geq \text{LIS}(Y_{\mathbf{a}} \circ \tau^{-1}). \quad (9)$$

Then, the lemma follows by Lemma 2.3. Because, let id denote the identity in S_k , we have

$$\begin{aligned} \text{LIS}(X_{\mathbf{a}} \circ \pi^{-1}) &= \text{LIS}(X_{\mathbf{a}} \circ \pi^{-1}, id) = \text{LIS}(X_{\mathbf{a}}, \pi), \\ \text{LIS}(Y_{\mathbf{a}} \circ \tau^{-1}) &= \text{LIS}(Y_{\mathbf{a}} \circ \tau^{-1}, id) = \text{LIS}(Y_{\mathbf{a}}, \tau). \end{aligned}$$

Suppose the set $\{\sigma, \sigma \circ (r, t)\}$, with $\sigma(r) < \sigma(t)$, is chosen in the first step. If the outcome in the second step is tail, we verify that $X_{\mathbf{a}} \circ \pi^{-1} = Y_{\mathbf{a}} \circ \tau^{-1}$. When $X = \sigma$, $Y = \sigma \circ (r, t)$, by (8), we have

$$\begin{aligned} X_{\mathbf{a}} \circ \pi^{-1} &= \sigma_{\mathbf{a}} \circ \pi^{-1}, \\ Y_{\mathbf{a}} \circ \tau^{-1} &= (\sigma \circ (r, t))_{\mathbf{a}} \circ \pi^{-1} \circ s_i \\ &= (\sigma \circ (r, t))_{\mathbf{a}} \circ (\pi^{-1}(i), \pi^{-1}(i+1)) \circ \pi^{-1} \\ &= (\sigma \circ (r, t) \circ (r, t))_{\mathbf{a}} \circ \pi^{-1} \\ &= \sigma_{\mathbf{a}} \circ \pi^{-1}. \end{aligned}$$

When $X = \sigma \circ (r, t)$, $Y = \sigma$, again by (8), we have

$$\begin{aligned} X_{\mathbf{a}} \circ \pi^{-1} &= (\sigma \circ (r, t))_{\mathbf{a}} \circ \pi^{-1} \\ &= \sigma_{\mathbf{a}} \circ (\pi^{-1}(i), \pi^{-1}(i+1)) \circ \pi^{-1} \\ &= \sigma_{\mathbf{a}} \circ \pi^{-1} \circ s_i, \\ Y_{\mathbf{a}} \circ \tau^{-1} &= \sigma_{\mathbf{a}} \circ \pi^{-1} \circ s_i. \end{aligned}$$

If the outcome in the second step is head, we have $X_{\mathbf{a}} \circ \pi^{-1} = \sigma_{\mathbf{a}} \circ \pi^{-1}$ and $Y_{\mathbf{a}} \circ \tau^{-1} = \sigma_{\mathbf{a}} \circ \pi^{-1} \circ s_i$. Since $\sigma(r) < \sigma(t)$, i.e., $\sigma(a_{\pi^{-1}(i)}) < \sigma(a_{\pi^{-1}(i+1)})$, we have $\sigma_{\mathbf{a}} \circ \pi^{-1}(i) < \sigma_{\mathbf{a}} \circ \pi^{-1}(i+1)$. Hence $Y_{\mathbf{a}} \circ \tau^{-1}$ covers $X_{\mathbf{a}} \circ \pi^{-1}$ in (S_k, \leq_R) . (9) follows. \square

Remark. A special case of Lemma 3.6 is when $k = n$, in which the only choice for \mathbf{a} is the vector $(1, 2, 3, \dots, n)$ whence $X_{\mathbf{a}} = X$, $Y_{\mathbf{a}} = Y$.

We can prove a similar result for the case when $q \geq 1$. Here we use the following property of Mallows permutation (cf. Lemma 2.2 in [3]).

Proposition 3.7. For any $n \geq 1$ and $q > 0$, if $\pi \sim \mu_{n,q}$ then $\pi^r \sim \mu_{n,1/q}$ and $\pi^{-1} \sim \mu_{n,q}$.

Lemma 3.8. Given $\pi, \tau \in S_k$ with $\pi \leq_L \tau$, for any $n \geq k$, $q \geq 1$ and $a_1 < \dots < a_k$ with $a_i \in [n]$, there exists a coupling (X, Y) such that $X \sim \mu_{n,q}$, $Y \sim \mu_{n,q}$ and

$$\text{LIS}(X_{\mathbf{a}}, \pi) \leq \text{LIS}(Y_{\mathbf{a}}, \tau).$$

Here $\mathbf{a} = (a_1, a_2, \dots, a_k)$.

Proof. Given $\pi \in S_n$, recall that π^r denote the reversal of π . For any $\pi \in S_n$, we have $\text{Inv}(\pi^r) = \{(i, j) : 1 \leq i < j \leq n \text{ and } (n+1-j, n+1-i) \notin \text{Inv}(\pi)\}$. Hence, $\pi \leq_L \tau$ implies $\pi^r \geq_L \tau^r$. By Lemma 3.6, there exists a coupling (U, V) such that $U \sim \mu_{n,1/q}$, $V \sim \mu_{n,1/q}$ and

$$\text{LIS}(U_{\mathbf{a}'}, \pi^r) \leq \text{LIS}(V_{\mathbf{a}'}, \tau^r).$$

Here $\mathbf{a}' = (a'_1, a'_2, \dots, a'_k)$ with $a'_i = n+1 - a_{k+1-i}$.

Define $(X, Y) := (U^r, V^r)$. By Proposition 3.7, $X \sim \mu_{n,q}$, $Y \sim \mu_{n,q}$. Moreover, we have

$$\begin{aligned} \text{LIS}(X_{\mathbf{a}}, \pi) &= \text{LIS}((X_{\mathbf{a}})^r, \pi^r) = \text{LIS}((X^r)_{\mathbf{a}'}, \pi^r) = \text{LIS}(U_{\mathbf{a}'}, \pi^r) \\ &\leq \text{LIS}(V_{\mathbf{a}'}, \tau^r) = \text{LIS}((Y^r)_{\mathbf{a}'}, \tau^r) = \text{LIS}((Y_{\mathbf{a}})^r, \tau^r) \\ &= \text{LIS}(Y_{\mathbf{a}}, \tau). \end{aligned}$$

\square

Lemma 3.9. *Given $\mathbf{a} = (a_1, a_2, \dots, a_k)$, where $a_1 < \dots < a_k$ and $a_i \in [n]$, for any $0 < q \leq 1$ and any distribution ν on S_k , there exists a coupling (X, Y, Z) such that the following holds,*

- (a) X and Y are independent.
- (b) $X \sim \mu_{n,q}$, $Y \sim \nu$ and $Z \sim \mu_{n,q}$.
- (c) $\text{LIS}(X_{\mathbf{a}}, Y) \leq \text{LIS}(Z_{\mathbf{a}})$.

Proof. Let id_k denote the identity in S_k . By the definition of weak bruhat order, for any $\xi \in S_k$, we have $id_k \leq_L \xi$. Hence, given $\xi \in S_k$, by Lemma 3.6, there exists a coupling (U, V) such that $U \sim \mu_{n,q}$, $V \sim \mu_{n,q}$ and $\text{LIS}(U_{\mathbf{a}}, \xi) \leq \text{LIS}(V_{\mathbf{a}}, id_k) = \text{LIS}(V_{\mathbf{a}})$. Then we construct the coupling (X, Y, Z) as follows

- Sample Y according to the distribution ν .
- Conditioned on $Y = \xi$, (X, Z) has the same distribution as (U, V) defined above.

First, we point out that X and Y are independent. Since whatever value Y takes, the conditional distribution of X is $\mu_{n,q}$. Moreover, it can be seen that X , Y and Z have the right marginal distribution. Finally, (c) holds by the construction of the coupling. \square

We can prove a similar result for the case when $q \geq 1$.

Lemma 3.10. *Given $\mathbf{a} = (a_1, a_2, \dots, a_k)$, where $a_1 < \dots < a_k$ and $a_i \in [n]$, for any $q \geq 1$ and any distribution ν on S_k , there exists a coupling (X, Y, Z) such that the following holds,*

- (a) X and Y are independent.
- (b) $X \sim \mu_{n,q}$, $Y \sim \nu$ and $Z \sim \mu_{n,q}$.
- (c) $\text{LIS}(X_{\mathbf{a}}, Y) \geq \text{LIS}(Z_{\mathbf{a}})$.

Proof. Let id_k denote the identity in S_k . By the definition of weak bruhat order, for any $\xi \in S_k$, we have $id_k \leq_L \xi$. Hence, given $\xi \in S_k$, by Lemma 3.8, there exists a coupling (U, V) such that $U \sim \mu_{n,q}$, $V \sim \mu_{n,q}$ and $\text{LIS}(U_{\mathbf{a}}, \xi) \geq \text{LIS}(V_{\mathbf{a}}, id_k) = \text{LIS}(V_{\mathbf{a}})$. Then we construct the coupling (X, Y, Z) as follows

- Sample Y according to the distribution ν .

- Conditioned on $Y = \xi$, (X, Z) has the same distribution as (U, V) defined above.

First, we point out that X and Y are independent. Since whatever value Y takes, the conditional distribution of X is $\mu_{n,q}$. Moreover, it can be seen that X , Y and Z have the right marginal distribution. Finally, (c) holds by the construction of the coupling. \square

Lemma 3.11. *Given $\mathbf{a} = (a_1, a_2, \dots, a_k)$, where $a_1 < \dots < a_k$ and $a_i \in [n]$. Define $\bar{\mathbf{a}} := \{n+1-a_k, n+1-a_{k-1}, \dots, n+1-a_1\}$. For any $0 < q \leq 1$ and any distribution ν on S_k , there exists a coupling (X, Y, Z) such that the following holds,*

- (a) X and Y are independent.
- (b) $X \sim \mu_{n,q}$, $Y \sim \nu$ and $Z \sim \mu_{n,1/q}$.
- (c) $\text{LIS}(X_{\mathbf{a}}, Y) \geq \text{LIS}(Z_{\bar{\mathbf{a}}})$.

Proof. Recall that π^r denotes the reversal of π . If $\pi \sim \nu$, we use ν^r to denote the distribution of π^r . Clearly, $\nu = (\nu^r)^r$. By Lemma 3.10, there exists a coupling (U, V, Z) such that

- U and V are independent.
- $U \sim \mu_{n,1/q}$, $V \sim \nu^r$ and $Z \sim \mu_{n,1/q}$.
- $\text{LIS}(U_{\bar{\mathbf{a}}}, V) \geq \text{LIS}(Z_{\bar{\mathbf{a}}})$.

Define $X := U^r$ and $Y := V^r$. We have

$$\begin{aligned}
\text{LIS}(U_{\bar{\mathbf{a}}}, V) &= \text{LIS}\left(\{(U_{\bar{\mathbf{a}}}(i), V(i))\}_{i \in [k]}\right) \\
&= \text{LIS}\left(\{((U_{\bar{\mathbf{a}}})^r(i), V^r(i))\}_{i \in [k]}\right) \\
&= \text{LIS}\left(\{((U^r)_{\mathbf{a}}(i), V^r(i))\}_{i \in [k]}\right) \\
&= \text{LIS}\left(\{(X_{\mathbf{a}}(i), Y(i))\}_{i \in [k]}\right) \\
&= \text{LIS}(X_{\mathbf{a}}, Y),
\end{aligned}$$

The lemma follows. \square

4 Proof of Theorem 1.4

We start this section by introducing the following two lemmas which can be seen as generalizations of Corollary 4.3 in [19]. That result shows that the LIS of a Mallows distributed permutation scaled by $n^{-1/2}$ can be bounded within a multiplicative interval of $e^{|\beta|}$ around 2. We postpone the proofs of these two lemmas to the end of this paper. For any positive integer n and $m \in [n]$, define

$$Q(n, m) := \{(b_1, b_2, \dots, b_m) : b_i \in [n] \text{ and } b_i < b_{i+1} \text{ for all } i\}.$$

Lemma 4.1. *Suppose that $\{q_n\}_{n=1}^\infty$ is a sequence such that $q_n \geq 1$ and $\liminf_{n \rightarrow \infty} n(1 - q_n) = \beta$, with $\beta \in \mathbb{R}$. For any sequence $\{k_n\}_{n=1}^\infty$ such that $k_n \in [n]$ and $\lim_{n \rightarrow \infty} k_n = \infty$, we have*

$$\lim_{n \rightarrow \infty} \max_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} \notin (2e^{\frac{\beta}{2}} - \epsilon, 2 + \epsilon) \right) = 0$$

for any $\epsilon > 0$.

Lemma 4.2. *Suppose that $\{q_n\}_{n=1}^\infty$ is a sequence such that $0 < q_n \leq 1$ and $\limsup_{n \rightarrow \infty} n(1 - q_n) = \beta < \ln 2$. For any sequence $\{k_n\}_{n=1}^\infty$ such that $k_n \in [n]$ and $\lim_{n \rightarrow \infty} k_n = \infty$, we have*

$$\lim_{n \rightarrow \infty} \max_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} \notin (2 - \epsilon, 2e^{\frac{\beta}{2}} + \epsilon) \right) = 0$$

for any $\epsilon > 0$.

Next, we introduce the following way to sample a permutation according to $\mu_{n, q}$ which will be used in the proofs. Given $\mathbf{c} = \{c_1, c_2, \dots, c_m\}$, where $c_i \in \mathbb{Z}^+$ and $\sum_{i=1}^m c_i = n$, define

$$d_0 := 0, \quad d_k := \sum_{i=1}^k c_i \quad \forall k \in [m],$$

$$A(\mathbf{c}) := \{(A_1, A_2, \dots, A_m) : \{A_i\}_{i \in [m]} \text{ is a partition of } [n], |A_i| = c_i\}.$$

Given $(A_1, \dots, A_m) \in A(\mathbf{c})$, define the inversion number of (A_1, \dots, A_m) as follows,

$$l((A_1, \dots, A_m)) :=$$

$$|\{(x, y) : x > y \text{ and there exists } i < j \text{ such that } x \in A_i, y \in A_j\}|.$$

Let \mathbf{a}_i be the vector which consists of the numbers in A_i in increasing order. There exists a bijection $f_{\mathbf{c}}$ between S_n and $A(\mathbf{c}) \times S_{c_1} \times S_{c_2} \times \cdots \times S_{c_m}$ such that, for any $\pi \in S_n$, $f_{\mathbf{c}}(\pi) = ((A_1, A_2, \dots, A_m), \tau_1, \tau_2, \dots, \tau_m)$ if and only if

$$\{\pi(j) : j \in A_i\} = \{d_{i-1} + 1, d_{i-1} + 2, \dots, d_i\}, \quad \pi_{\mathbf{a}_i} = \tau_i, \quad \forall i \in [m].$$

From the definition above, it is not hard to see that the following relation holds,

$$l(\pi) = l((A_1, A_2, \dots, A_m)) + \sum_{i=1}^m l(\tau_i). \quad (10)$$

Define the random variable $X_{\mathbf{c}}$ which takes value in $A(\mathbf{c})$ such that

$$\mathbb{P}(X_{\mathbf{c}} = (A_1, A_2, \dots, A_m)) \propto q^{l((A_1, A_2, \dots, A_m))}.$$

Independent of $X_{\mathbf{c}}$, let Y_1, Y_2, \dots, Y_m be independent random variables such that, for any $i \in [m]$, $Y_i \sim \mu_{c_i, q}$. Define $Z := f_{\mathbf{c}}^{-1}(X_{\mathbf{c}}, Y_1, Y_2, \dots, Y_m)$. By (10), we have $Z \sim \mu_{n, q}$, since

$$\mathbb{P}(Z = \pi) \propto q^{l(\pi)}.$$

As our last step in preparation for the proof of Lemma 4.4, we introduce the following elementary result in analysis.

Lemma 4.3. *Suppose $\{B_i\}_{i=1}^{\infty}$ is a partition of \mathbb{N} , i.e. $\cup_{i=1}^{\infty} B_i = \mathbb{N}$ and $B_i \cap B_j = \emptyset$, $\forall i \neq j$. Moreover, each B_i is a finite nonempty set. Given a sequence $\{x_i\}_{i=1}^{\infty}$, if $\lim_{n \rightarrow \infty} x_{b_n} = a$, for any sequence $\{b_i\}_{i=1}^{\infty}$ with $b_i \in B_i$, then we have $\lim_{n \rightarrow \infty} x_n = a$.*

Proof. We prove the lemma by contradiction. Suppose $\lim_{n \rightarrow \infty} x_n = a$ does not hold. Then there exists $\epsilon > 0$ and a subsequence $\{x_{n_j}\}_{j=1}^{\infty}$ such that $x_{n_j} \notin (a - \epsilon, a + \epsilon)$ for all j . Since each B_i is a finite set, without loss of generality, we may assume that each B_i contains at most one n_j . Then, we can construct a sequence $\{b_i\}_{i=1}^{\infty}$ with $b_i \in B_i$, such that $x_{b_i} \notin (a - \epsilon, a + \epsilon)$ infinitely often. Specifically, we define the sequence $\{b_i\}_{i=1}^{\infty}$ as follows. For each i , if there exists an $n_j \in B_i$, let $b_i = n_j$, otherwise, let b_i be an arbitrary number in B_i . Thus, we get the contradiction. \square

For any $\pi, \tau \in S_n$, define $\mathbf{z}(\pi, \tau) := \{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i \in [n]}$. Let $l_R(\pi, \tau)$ denote the length of the longest increasing subsequence of $\mathbf{z}(\pi, \tau)$ within R . The following lemma addresses the size of the LIS in a small rectangle and this result will be the most crucial building block used to show both the upper and lower bounds of Theorem 1.4.

Lemma 4.4. *Let $R = (x_1, x_2] \times (y_1, y_2] \subset [0, 1] \times [0, 1]$. Under the same conditions as in Lemma 1.3, if $|\Delta x| |\beta| < \ln 2$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\frac{l_R(\pi, \tau)}{\sqrt{n\rho(R)}} \in \left(2e^{-\Delta x |\beta|/2} - \epsilon, 2e^{\Delta x |\beta|/2} + \epsilon \right) \right) = 1, \quad (11)$$

for any $\epsilon > 0$, where $\rho(R) := \iint_R \rho(x, y) dx dy$ and $\Delta x := x_2 - x_1$.

Proof. To simplify the proof, we divide the lemma into the following three cases:

Case 1: $\beta > 0$ or $\beta = 0$ and $q_n \leq 1$ when n is sufficiently large.

Case 2: $\beta < 0$ or $\beta = 0$ and $q_n \geq 1$ when n is sufficiently large.

Case 3: $\beta = 0$.

Firstly, **Case 3** follows from **Case 1** and **Case 2** because if $\lim_{n \rightarrow \infty} n(1 - q_n) = 0$, we can divide the sequence $\{q_n\}_{n=1}^\infty$ into two disjoint subsequences such that one of them falls into **Case 1** and the other falls into **Case 2**.

Next we argue that **Case 2** follows from **Case 1**. If $\pi \sim \mu_{n,q}$, by Proposition 3.7, we have $\pi^r \sim \mu_{n,1/q}$. Trivially, for any $\pi, \tau \in S_n$, we have $\mathbf{z}(\pi, \tau) = \mathbf{z}(\pi^r, \tau^r)$. Since $\lim_{n \rightarrow \infty} n(1 - q_n) = \beta \in \mathbb{R}$, we have $\lim_{n \rightarrow \infty} q_n = 1$. Hence,

$$\lim_{n \rightarrow \infty} n(1 - 1/q_n) = \lim_{n \rightarrow \infty} n(q_n - 1)/q_n = -\beta.$$

Therefore, **Case 2** follows from **Case 1** by considering the reversal of π and τ in (11). Specifically, if $\pi \sim \mu_{n,q_n}$ and $\tau \sim \mu_{n,q'_n}$, after reversing, we have $\pi^r \sim \mu_{n,1/q_n}$ and $\tau^r \sim \mu_{n,1/q'_n}$ and the n points induced by π and τ do not change, i.e., $\mathbf{z}(\pi, \tau) = \mathbf{z}(\pi^r, \tau^r)$.

To prove **Case 1**, in the following, we assume $x_1, y_1 > 0$ and $x_2, y_2 < 1$. The proofs for the cases when $x_1 = 0$ or $y_1 = 0$ or $x_2 = 1$ or $y_2 = 1$ are similar. Let $x_3 = y_3 = 1$. Given $n \in \mathbb{N}$, we will sample (π, τ) according to \mathbb{P}_n by the method introduced before Lemma 4.3. Define

$$d_{n,i} := \lfloor nx_i \rfloor, \quad c_{n,i} := d_{n,i} - d_{n,i-1}, \quad \text{for } i = 1, 2, 3,$$

$$d'_{n,i} := \lfloor ny_i \rfloor, \quad c'_{n,i} := d'_{n,i} - d'_{n,i-1}, \quad \text{for } i = 1, 2, 3.$$

Here we assume that $d_{n,0} = d'_{n,0} = 0$. Then, it is trivial that

$$\begin{aligned} d_{n,i} &= |\{j \in [n] : \frac{j}{n} \in (0, x_i)\}|, & c_{n,2} &= |\{j \in [n] : \frac{j}{n} \in (x_1, x_2)\}|, \\ d'_{n,i} &= |\{j \in [n] : \frac{j}{n} \in (0, y_i)\}|, & c'_{n,2} &= |\{j \in [n] : \frac{j}{n} \in (y_1, y_2)\}|. \end{aligned}$$

Since $\lim_{n \rightarrow \infty} \frac{\lfloor nx \rfloor}{n} = x, \forall x \in \mathbb{R}$, it follows that $\lim_{n \rightarrow \infty} \frac{d_{n,i}}{n} = x_i$. Hence

$$\lim_{n \rightarrow \infty} \frac{c_{n,2}}{n} = x_2 - x_1 = \Delta x. \quad (12)$$

Next, for any nonnegative integer i , define $B_i := \{n \in \mathbb{N} : c_{n,2} = i\}$. Clearly, $\{B_i\}_{i=0}^\infty$ thus defined is a partition of \mathbb{N} and we show that each B_i is a nonempty finite set. Since, by (12), $\lim_{n \rightarrow \infty} c_{n,2} = \infty$, we conclude that each B_i is a finite set. From the definition of $d_{n,i}$, it is easily seen that the sequence $\{d_{n,1}\}$ is nondecreasing and the increment of consecutive terms is either 0 or 1. The same is true for the sequence $\{d_{n,2}\}$. Hence, we have

$$|c_{n+1,2} - c_{n,2}| = |d_{n+1,2} - d_{n,2} - (d_{n+1,1} - d_{n,1})| \leq 1.$$

Since $c_{1,2} \in B_0$ and $\lim_{n \rightarrow \infty} c_{n,2} = \infty$, the inequality above guarantees that each B_i is nonempty.

Next, define $\mathbf{c}_n = (c_{n,1}, c_{n,2}, c_{n,3})$ and $\mathbf{c}'_n = (c'_{n,1}, c'_{n,2}, c'_{n,3})$. Define $X_{\mathbf{c}_n}$ which takes values in $A(\mathbf{c}_n)$ such that

$$\mathbb{P}(X_{\mathbf{c}_n} = (A_1, A_2, A_3)) \propto q_n^{l((A_1, A_2, A_3))}, \quad \forall (A_1, A_2, A_3) \in A(\mathbf{c}_n).$$

Independently, define three independent random variables $Y_{n,1}, Y_{n,2}, Y_{n,3}$ such that $Y_{n,i} \sim \mu_{c_{n,i}, q_n}$. Independent of all the variables defined above, define $X_{\mathbf{c}'_n}$ and $Y'_{n,1}, Y'_{n,2}, Y'_{n,3}$ in the same fashion. That is, $X_{\mathbf{c}'_n}$ takes value in $A(\mathbf{c}'_n)$ with

$$\mathbb{P}(X_{\mathbf{c}'_n} = (A'_1, A'_2, A'_3)) \propto (q'_n)^{l((A'_1, A'_2, A'_3))}, \quad \forall (A'_1, A'_2, A'_3) \in A(\mathbf{c}'_n)$$

and $Y'_{n,1}, Y'_{n,2}, Y'_{n,3}$ are three independent random variables with $Y'_{n,i} \sim \mu_{c'_{n,i}, q'_n}$. Define

$$\pi := f_{\mathbf{c}_n}^{-1}(X_{\mathbf{c}_n}, Y_{n,1}, Y_{n,2}, Y_{n,3}), \quad \tau := f_{\mathbf{c}'_n}^{-1}(X_{\mathbf{c}'_n}, Y'_{n,1}, Y'_{n,2}, Y'_{n,3}).$$

From the discussion before Lemma 4.3, it follows that (π, τ) thus defined has distribution \mathbb{P}_n . Moreover, given $X_{\mathbf{c}_n} = (A_1, A_2, A_3)$ and $X_{\mathbf{c}'_n} = (A'_1, A'_2, A'_3)$, we have

$$A_2 = \left\{ i \in [n] : \frac{\pi(i)}{n} \in (x_1, x_2] \right\}, \quad A'_2 = \left\{ i \in [n] : \frac{\tau(i)}{n} \in (y_1, y_2] \right\}.$$

Hence, we have

$$A_2 \cap A'_2 = \left\{ i \in [n] : \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n} \right) \in R \right\}. \quad (13)$$

Define $M = |\mathbf{z}(\pi, \tau) \cap R|$, i. e. M denotes the number of points $\{(\frac{\pi(i)}{n}, \frac{\tau(i)}{n})\}_{i=1}^n$ within R . Then, by (13), we have $M = |A_2 \cap A'_2|$. Hence, M only depends on the values of $X_{\mathbf{c}_n}$ and $X_{\mathbf{c}'_n}$ and is independent of $\cup_{i \in [3]} \{Y_{n,i}, Y'_{n,i}\}$. Next, we point out that, conditioning on $X_{\mathbf{c}_n} = (A_1, A_2, A_3)$ and $X_{\mathbf{c}'_n} = (A'_1, A'_2, A'_3)$, $l_R(\pi, \tau)$ is determined by $Y_{n,2}$ and $Y'_{n,2}$. To see this, we first define a new function I as follows, given any finite set $A \subset \mathbb{Z}$ and any $a \in A$, define $I(A, a) := k$ if a is the k -th smallest number in A . Suppose $A_2 \cap A'_2 = \{a_j\}_{j \in [M]}$ with $a_1 < a_2 < \dots < a_M$. Define $\mathbf{b} \in Q(c_{n,2}, M)$ and $\mathbf{b}' \in Q(c'_{n,2}, M)$ by

$$\begin{aligned} \mathbf{b} &:= (I(A_2, a_1), I(A_2, a_2), \dots, I(A_2, a_M)), \\ \mathbf{b}' &:= (I(A'_2, a_1), I(A'_2, a_2), \dots, I(A'_2, a_M)). \end{aligned} \quad (14)$$

Note that \mathbf{b} and \mathbf{b}' are determined by A_2 and A'_2 . Then, we have

$$l_R(\pi, \tau) = \text{LIS}((Y_{n,2})_{\mathbf{b}}, (Y'_{n,2})_{\mathbf{b}'}). \quad (15)$$

Because, conditioning on $X_{\mathbf{c}_n} = (A_1, A_2, A_3)$, we know that $\{\pi(i) : i \in A_2\} = \{d_{n,1} + 1, d_{n,1} + 2, \dots, d_{n,2}\}$. And the value of $Y_{n,2}$ determines the relative ordering of $\pi(i)$ for those $i \in A_2$. Similarly, the value of $Y'_{n,2}$ determines the relative ordering of $\tau(i)$ for those $i \in A'_2$.

Now we are in the position to prove (11) for **Case 1**. From the discussion above and Lemma 4.3, it suffices to show that, for any sequence $\{s_n\}_{n=1}^\infty$ with $s_n \in B_n$, i.e., when $c_{s_n,2} = n$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_{s_n} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} \in \left(2e^{-\Delta x \beta/2} - \epsilon, 2e^{\Delta x \beta/2} + \epsilon \right) \right) = 1, \quad (16)$$

for any $\epsilon > 0$. Note that by the definition of \mathbb{P}_{s_n} in Lemma 1.3, π and τ above are of size s_n with $\pi \sim \mu_{s_n, q_{s_n}}$, $\tau \sim \mu_{s_n, q'_{s_n}}$.

We separate the proof of (16) into two parts. Specifically, we need to show that

$$\lim_{n \rightarrow \infty} \mathbb{P}_{s_n} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} < 2e^{\Delta x \beta/2} + \epsilon \right) = 1, \quad (17)$$

and

$$\lim_{n \rightarrow \infty} \mathbb{P}_{s_n} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \epsilon \right) = 1, \quad (18)$$

for any $\epsilon > 0$.

Since $\{s_n\}_{n \geq 1}$ is a subsequence of $\{i\}_{i \geq 0}$, $\lim_{n \rightarrow \infty} s_n = \infty$. Hence, by (12) and the fact that $c_{s_n,2} = n$, we get

$$\lim_{n \rightarrow \infty} \frac{n}{s_n} = \lim_{n \rightarrow \infty} \frac{c_{s_n,2}}{s_n} = \Delta x.$$

Thus,

$$\lim_{n \rightarrow \infty} n(1 - q_{s_n}) = \lim_{n \rightarrow \infty} \frac{n}{s_n} s_n(1 - q_{s_n}) = \Delta x \beta < \ln 2. \quad (19)$$

To prove (17), for any $\epsilon > 0$, we can choose $\epsilon_1 > 0$ sufficiently small such that

$$(1 - \epsilon_1)(2e^{\Delta x \beta/2} + \epsilon) > 2e^{\Delta x \beta/2}. \quad (20)$$

For this fixed ϵ_1 , we can choose $\delta > 0$ such that

$$\sqrt{\frac{\rho(R)}{\rho(R) + \delta}} > 1 - \epsilon_1. \quad (21)$$

Given $n \in \mathbb{N}$, define $k_n = \lfloor s_n(\rho(R) + \delta) \rfloor$. Clearly, we have $\lim_{n \rightarrow \infty} k_n = \infty$. Moreover, under the conditions of **Case 1**, $q_n \leq 1$ for sufficiently large n . Hence, by Lemma 4.2, (19) and (20), there exists $N_1 > 0$ such that, for any $n > N_1$, we have

$$\min_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_{s_n}} \left(\eta \in S_n : \frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{k_n}} < (1 - \epsilon_1)(2e^{\Delta x \beta/2} + \epsilon) \right) > 1 - \epsilon. \quad (22)$$

Given $\mathbf{b} \in Q(n, k_n)$, for any \mathbf{b}' which is a subsequence of \mathbf{b} , we have $\text{LIS}(\eta_{\mathbf{b}}) \geq \text{LIS}(\eta_{\mathbf{b}'})$. Thus we can make (22) stronger as follows,

$$\min_{\mathbf{b} \in \bar{Q}(n, k_n)} \mu_{n, q_{s_n}} \left(\eta \in S_n : \frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{k_n}} < (1 - \epsilon_1)(2e^{\Delta x \beta/2} + \epsilon) \right) > 1 - \epsilon, \quad (23)$$

where $\bar{Q}(n, k_n) = \cup_{i \in [k_n]} Q(n, i)$. Since $\lim_{n \rightarrow \infty} s_n = \infty$, we have

$$\lim_{n \rightarrow \infty} s_n(1 - q_{s_n}) = \beta \quad \text{and} \quad \lim_{n \rightarrow \infty} s_n(1 - q'_{s_n}) = \gamma. \quad (24)$$

Hence, by Lemma 1.3, there exists $N_2 > 0$ such that, for any $n > N_2$, we have

$$\mathbb{P}_{s_n} \left(\frac{|\mathbf{z}(\pi, \tau) \cap R|}{s_n} \leq \rho(R) + \delta \right) > 1 - \epsilon. \quad (25)$$

In the following, let $E_n(A_2, A'_2)$ denote the event that the second entries of $X_{\mathbf{c}_{s_n}}$ and $X_{\mathbf{c}'_{s_n}}$ are A_2 and A'_2 respectively. Then, for any $n > \max(N_1, N_2)$, we have

$$\begin{aligned}
& \mathbb{P}_{s_n} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} < 2e^{\Delta x \beta/2} + \epsilon \right) \\
& \geq \sum_{|A_2 \cap A'_2| \leq k_n} \mathbb{P} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} < 2e^{\Delta x \beta/2} + \epsilon \mid E_n(A_2, A'_2) \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
& = \sum_{|A_2 \cap A'_2| \leq k_n} \mathbb{P} \left(\frac{\text{LIS}((Y_{s_n,2})_{\mathbf{b}}, (Y'_{s_n,2})_{\mathbf{b}'})}{\sqrt{s_n \rho(R)}} < 2e^{\Delta x \beta/2} + \epsilon \mid E_n(A_2, A'_2) \right) \\
& \quad \times \mathbb{P}(E_n(A_2, A'_2)) \\
& = \sum_{|A_2 \cap A'_2| \leq k_n} \mathbb{P} \left(\frac{\text{LIS}((Y_{s_n,2})_{\mathbf{b}}, (Y'_{s_n,2})_{\mathbf{b}'})}{\sqrt{s_n \rho(R)}} < 2e^{\Delta x \beta/2} + \epsilon \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
& \geq \sum_{|A_2 \cap A'_2| \leq k_n} \mu_{n, q_{s_n}} \left(\frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{s_n \rho(R)}} < 2e^{\Delta x \beta/2} + \epsilon \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
& = \sum_{|A_2 \cap A'_2| \leq k_n} \mu_{n, q_{s_n}} \left(\frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{s_n(\rho(R) + \delta)}} < \frac{\sqrt{\rho(R)}}{\sqrt{\rho(R) + \delta}} (2e^{\Delta x \beta/2} + \epsilon) \right) \\
& \quad \times \mathbb{P}(E_n(A_2, A'_2)) \\
& \geq \sum_{|A_2 \cap A'_2| \leq k_n} \mu_{n, q_{s_n}} \left(\frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{k_n}} < (1 - \epsilon_1)(2e^{\Delta x \beta/2} + \epsilon) \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
& \geq (1 - \epsilon) \times \sum_{|A_2 \cap A'_2| \leq k_n} \mathbb{P}(E_n(A_2, A'_2)) \\
& = (1 - \epsilon) \times \mathbb{P}_{s_n}(|\mathbf{z}(\pi, \tau) \cap R| \leq k_n) \\
& = (1 - \epsilon) \times \mathbb{P}_{s_n}(|\mathbf{z}(\pi, \tau) \cap R| \leq s_n(\rho(R) + \delta)) \\
& > (1 - \epsilon)^2.
\end{aligned}$$

Here \mathbb{P} denotes the probability space on which $(X_{\mathbf{c}_{s_n}}, Y_{s_n,1}, Y_{s_n,2}, Y_{s_n,3})$ and $(X_{\mathbf{c}'_{s_n}}, Y'_{s_n,1}, Y'_{s_n,2}, Y'_{s_n,3})$ are defined. The first equality follows by (15). The second equality follows by independence of $(X_{\mathbf{c}_{s_n}}, X_{\mathbf{c}'_{s_n}})$ and $(Y_{s_n,2}, Y'_{s_n,2})$. Note that \mathbf{b} and \mathbf{b}' are determined by A_2 and A'_2 as in (14). The second inequality follows by Lemma 3.9, since $Y_{s_n,2}$ and $Y'_{s_n,2}$ are independent with $Y_{s_n,2} \sim \mu_{n, q_{s_n}}$. The third inequality follows by (21) and the fact that $k_n = \lfloor s_n(\rho(R) + \delta) \rfloor \leq s_n(\rho(R) + \delta)$. The fourth inequality follows by (23) and the fact that the dimension of \mathbf{b} equals to $|A_2 \cap A'_2|$. The last inequality

follows by (25). Hence, (17) follows.

The proof of (18) is analogous to the proof of (17). First, by (19) and the fact that $\lim_{n \rightarrow \infty} q_n = 1$, we have

$$\lim_{n \rightarrow \infty} n(1 - 1/q_{s_n}) = \lim_{n \rightarrow \infty} \frac{n(q_{s_n} - 1)}{q_{s_n}} = -\Delta x \beta. \quad (26)$$

For any $\epsilon > 0$, we can choose $\epsilon_1 > 0$ sufficiently small such that

$$(1 + \epsilon_1)(2e^{-\Delta x \beta/2} - \epsilon) < 2e^{-\Delta x \beta/2}. \quad (27)$$

For this fixed ϵ_1 , we can choose $\delta > 0$ such that

$$\sqrt{\frac{\rho(R)}{\rho(R) - \delta}} < 1 + \epsilon_1. \quad (28)$$

Given $n \in \mathbb{N}$, define $k'_n = \lceil s_n(\rho(R) - \delta) \rceil$. Clearly, we have $\lim_{n \rightarrow \infty} k'_n = \infty$. Moreover, under conditions of **Case 1**, $1/q_n \geq 1$ for sufficiently large n . Hence, by Lemma 4.1, (26) and (27), there exist $N_3 > 0$ such that, for any $n > N_3$, we have

$$\min_{\mathbf{b} \in Q(n, k'_n)} \mu_{n, 1/q_{s_n}} \left(\eta \in S_n : \frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{k'_n}} > (1 + \epsilon_1)(2e^{-\Delta x \beta/2} - \epsilon) \right) > 1 - \epsilon. \quad (29)$$

Given $\mathbf{b} \in Q(n, k'_n)$, for any \mathbf{b}' such that \mathbf{b} is a subsequence of \mathbf{b}' , we have $\text{LIS}(\eta_{\mathbf{b}}) \leq \text{LIS}(\eta_{\mathbf{b}'})$. Thus we can make (29) stronger as follows,

$$\min_{\mathbf{b} \in \hat{Q}(n, k'_n)} \mu_{n, 1/q_{s_n}} \left(\eta \in S_n : \frac{\text{LIS}(\eta_{\mathbf{b}})}{\sqrt{k'_n}} > (1 + \epsilon_1)(2e^{-\Delta x \beta/2} - \epsilon) \right) > 1 - \epsilon, \quad (30)$$

where $\hat{Q}(n, k'_n) = \cup_{k'_n \leq i \leq n} Q(n, i)$.

By (24) and Lemma 1.3, there exists $N_4 > 0$ such that, for any $n > N_4$, we have

$$\mathbb{P}_{s_n} \left(\frac{|\mathbf{z}(\pi, \tau) \cap R|}{s_n} \geq \rho(R) - \delta \right) > 1 - \epsilon. \quad (31)$$

Again, let $E_n(A_2, A'_2)$ denote the event that the second entries of $X_{\mathbf{c}_{s_n}}$ and $X_{\mathbf{c}'_{s_n}}$ are A_2 and A'_2 respectively. Then, for any $n > \max(N_3, N_4)$, we have

$$\mathbb{P}_{s_n} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \epsilon \right)$$

$$\begin{aligned}
&\geq \sum_{|A_2 \cap A'_2| \geq k_n} \mathbb{P} \left(\frac{l_R(\pi, \tau)}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \epsilon \mid E_n(A_2, A'_2) \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
&= \sum_{|A_2 \cap A'_2| \geq k'_n} \mathbb{P} \left(\frac{\text{LIS}((Y_{s_n,2})_{\mathbf{b}}, (Y'_{s_n,2})_{\mathbf{b}'})}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \epsilon \mid E_n(A_2, A'_2) \right) \\
&\quad \times \mathbb{P}(E_n(A_2, A'_2)) \\
&= \sum_{|A_2 \cap A'_2| \geq k'_n} \mathbb{P} \left(\frac{\text{LIS}((Y_{s_n,2})_{\mathbf{b}}, (Y'_{s_n,2})_{\mathbf{b}'})}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \epsilon \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
&\geq \sum_{|A_2 \cap A'_2| \geq k'_n} \mu_{n,1/q_{s_n}} \left(\frac{\text{LIS}(\eta_{\bar{\mathbf{b}}})}{\sqrt{s_n \rho(R)}} > 2e^{-\Delta x \beta/2} - \epsilon \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
&= \sum_{|A_2 \cap A'_2| \geq k'_n} \mu_{n,1/q_{s_n}} \left(\frac{\text{LIS}(\eta_{\bar{\mathbf{b}}})}{\sqrt{s_n(\rho(R) - \delta)}} > \frac{\sqrt{\rho(R)}}{\sqrt{\rho(R) - \delta}} (2e^{-\Delta x \beta/2} - \epsilon) \right) \\
&\quad \times \mathbb{P}(E_n(A_2, A'_2)) \\
&\geq \sum_{|A_2 \cap A'_2| \geq k'_n} \mu_{n,1/q_{s_n}} \left(\frac{\text{LIS}(\eta_{\bar{\mathbf{b}}})}{\sqrt{k'_n}} > (1 + \epsilon_1)(2e^{-\Delta x \beta/2} - \epsilon) \right) \times \mathbb{P}(E_n(A_2, A'_2)) \\
&\geq (1 - \epsilon) \times \sum_{|A_2 \cap A'_2| \geq k'_n} \mathbb{P}(E_n(A_2, A'_2)) \\
&= (1 - \epsilon) \times \mathbb{P}_{s_n}(|\mathbf{z}(\pi, \tau) \cap R| \geq k'_n) \\
&= (1 - \epsilon) \times \mathbb{P}_{s_n}(|\mathbf{z}(\pi, \tau) \cap R| \geq s_n(\rho(R) - \delta)) \\
&> (1 - \epsilon)^2.
\end{aligned}$$

Here \mathbb{P} denotes the probability space on which $(X_{\mathbf{c}_{s_n}}, Y_{s_n,1}, Y_{s_n,2}, Y_{s_n,3})$ and $(X_{\mathbf{c}'_{s_n}}, Y'_{s_n,1}, Y'_{s_n,2}, Y'_{s_n,3})$ are defined. The first equality follows by (15). The second equality follows by independence of $(X_{\mathbf{c}_{s_n}}, X_{\mathbf{c}'_{s_n}})$ and $(Y_{s_n,2}, Y'_{s_n,2})$. The second inequality follows by Lemma 3.11, since $Y_{s_n,2}$ and $Y'_{s_n,2}$ are independent with $Y_{s_n,2} \sim \mu_{n,q_{s_n}}$. The third inequality follows by (28) and the fact that $k'_n = \lceil s_n(\rho(R) - \delta) \rceil \geq s_n(\rho(R) - \delta)$. The fourth inequality follows by (30) and the fact that $\bar{\mathbf{b}}$ has the same dimension as of \mathbf{b} which equals to $|A_2 \cap A'_2|$. The last inequality follows by (31). Hence, (18) follows and this completes the proof of Lemma 4.4. \square

The following lemma establishes certain degree of smoothness of the densities u and ρ defined in Lemma 1.3.

Lemma 4.5. *The density functions $u(x, y, \beta)$ defined in (3) and $\rho(x, y)$ defined in (2) satisfy the following,*

$$(a) \quad e^{-|\beta|} \leq u(x, y, \beta) \leq e^{|\beta|}, \quad e^{-|\beta|-|\gamma|} \leq \rho(x, y) \leq e^{|\beta|+|\gamma|},$$

$$(b) \quad u(x, y, \beta) \in C_b^1, \quad \rho(x, y) \in C_b^1,$$

$$(c) \quad \max \left(\left| \frac{\partial u}{\partial x} \right|, \left| \frac{\partial u}{\partial y} \right| \right) \leq |\beta| e^{|\beta|},$$

$$(d) \quad \max \left(\left| \frac{\partial \rho}{\partial x} \right|, \left| \frac{\partial \rho}{\partial y} \right| \right) \leq (|\beta| + |\gamma|) e^{|\beta|+|\gamma|},$$

where $(x, y) \in [0, 1] \times [0, 1]$.

Proof. First we show that $e^{-|\beta|} \leq u(x, y, \beta) \leq e^{|\beta|}$ for any $0 \leq x, y \leq 1$. Here we assume $\beta > 0$. The proof for the case when $\beta < 0$ is similar. By (3), we have

$$\begin{aligned} u(x, y, \beta) &= \frac{(\beta/2) \sinh(\beta/2)}{(e^{\beta/4} \cosh(\beta[x-y]/2) - e^{-\beta/4} \cosh(\beta[x+y-1]/2))^2} \\ &= \frac{\beta(e^\beta - 1)}{(2e^{\beta/2} \cosh(\beta[x-y]/2) - 2 \cosh(\beta[x+y-1]/2))^2}. \end{aligned} \quad (32)$$

Since $-1 \leq x - y \leq 1$ and $-1 \leq x + y - 1 \leq 1$, we have

$$2e^{\beta/2} \leq 2e^{\beta/2} \cosh(\beta[x-y]/2) \leq e^\beta + 1, \quad (33)$$

$$2 \leq 2 \cosh(\beta[x+y-1]/2) \leq e^{\beta/2} + e^{-\beta/2}. \quad (34)$$

Since $e^{\beta/2} + e^{-\beta/2} < 2e^{\beta/2}$, from (33) and (34), we have

$$e^{\beta/2} - e^{-\beta/2} \leq 2e^{\beta/2} \cosh(\beta[x-y]/2) - 2 \cosh(\beta[x+y-1]/2) \leq e^\beta - 1. \quad (35)$$

By (32) and (35), it follows that

$$\frac{\beta}{e^\beta - 1} \leq u(x, y, \beta) \leq \frac{\beta(e^\beta - 1)}{(e^{\beta/2} - e^{-\beta/2})^2}. \quad (36)$$

It is easily verified that

$$\frac{\beta}{e^\beta - 1} \geq e^{-\beta} \iff e^{-\beta} \geq 1 - \beta, \quad (37)$$

$$\frac{\beta(e^\beta - 1)}{(e^{\beta/2} - e^{-\beta/2})^2} \leq e^\beta \iff (e^\beta - 1)(e^\beta - 1 - \beta) \geq 0. \quad (38)$$

By the inequality $e^x \geq 1 + x$, the right-hand side of (37) and (38) hold. It follows from (36) and the left-hand side of (37) and (38) that

$$e^{-\beta} \leq u(x, y, \beta) \leq e^\beta, \quad \forall 0 \leq x, y \leq 1.$$

By the definition of $\rho(x, y)$, it follows trivially that

$$e^{-|\beta|+|\gamma|} \leq \rho(x, y) \leq e^{|\beta|+|\gamma|}, \quad \forall 0 \leq x, y \leq 1.$$

In [21], Starr shows that $\frac{\partial^2 \ln u(x, y, \beta)}{\partial x \partial y} = 2\beta u(x, y, \beta)$. Thus

$$\int_0^x u(t, y, \beta) dt = \frac{1}{2\beta} \left(\frac{\partial \ln u(x, y, \beta)}{\partial y} - \frac{\partial \ln u(0, y, \beta)}{\partial y} \right). \quad (39)$$

By direct calculation, we have $u(1, y, \beta) = \frac{\beta e^{\beta y}}{e^{\beta} - 1}$, $u(0, y, \beta) = \frac{\beta e^{-\beta y}}{1 - e^{-\beta}}$. Therefore, we get $\frac{\partial \ln u(1, y, \beta)}{\partial y} = \beta$ and $\frac{\partial \ln u(0, y, \beta)}{\partial y} = -\beta$. By (39), it follows that

$$\frac{\partial u(x, y, \beta)}{\partial y} = 2\beta u(x, y, \beta) \left(\int_0^x u(t, y, \beta) dt - \frac{1}{2} \right), \quad (40)$$

and

$$\int_0^x u(t, y, \beta) dt \leq \int_0^1 u(t, y, \beta) dt = 1. \quad (41)$$

From (40) and (41), we get

$$\left| \frac{\partial u}{\partial y} \right| \leq |\beta| u(x, y, \beta) \leq |\beta| e^{|\beta|}. \quad (42)$$

Since $u(x, y, \beta)$ is uniformly continuous on $[0, 1] \times [0, 1]$, $\int_0^x u(t, y, \beta) dt$ is also continuous on $[0, 1] \times [0, 1]$. Hence, by (40), $\frac{\partial u}{\partial y}$ is bounded and continuous on $[0, 1] \times [0, 1]$. Similar argument can be made for $\frac{\partial u}{\partial x}$. Thus we have shown that $u(x, y, \beta) \in C_b^1$ and

$$\max \left(\left| \frac{\partial u}{\partial x} \right|, \left| \frac{\partial u}{\partial y} \right| \right) \leq |\beta| e^{|\beta|}.$$

Next, since $\left| \frac{\partial u(x, t, \beta)}{\partial x} \cdot u(t, y, \gamma) \right| \leq |\beta| e^{|\beta|+|\gamma|}$ for any $0 \leq x, y, t \leq 1$, by dominated convergence theorem, we have

$$\frac{\partial \rho(x, y)}{\partial x} = \frac{\partial}{\partial x} \left(\int_0^1 u(x, t, \beta) u(t, y, \gamma) dt \right) = \int_0^1 \frac{\partial u(x, t, \beta)}{\partial x} u(t, y, \gamma) dt. \quad (43)$$

Hence, $\left| \frac{\partial \rho}{\partial x} \right| \leq |\beta| e^{|\beta|+|\gamma|}$. Moreover, $\frac{\partial u(x, t, \beta)}{\partial x} \cdot u(t, y, \gamma)$ as a function of x, y, t is uniformly continuous on $[0, 1] \times [0, 1] \times [0, 1]$. Thus, by (43), $\frac{\partial \rho}{\partial x}$ is continuous on $[0, 1] \times [0, 1]$. By a similar argument, it can be shown that $\frac{\partial \rho}{\partial y}$ is continuous on $[0, 1] \times [0, 1]$, and $\left| \frac{\partial \rho}{\partial y} \right| \leq |\gamma| e^{|\beta|+|\gamma|}$. Therefore, $\rho(x, y) \in C_b^1$ and

$$\max \left(\left| \frac{\partial \rho}{\partial x} \right|, \left| \frac{\partial \rho}{\partial y} \right| \right) \leq (|\beta| + |\gamma|) e^{|\beta|+|\gamma|}.$$

□

The next lemma shows that for any non-decreasing curve in the unit square, in a strip of small width around it, with probability going to 1, there exists an increasing subsequence whose length can be bounded from below. The proof of Lemma 4.7 uses similar arguments as in the proof of Lemma 8 in [9]. Before stating the lemma, we need the following notation.

Definition 4.6. Let B_{\nearrow} be the set of nondecreasing, right continuous functions $\phi : [0, 1] \rightarrow [0, 1]$. For $\phi \in B_{\nearrow}$, we have $\phi(x) = \int_0^x \dot{\phi}(t) dt + \phi_s(x)$, where ϕ_s is singular and has a zero derivative almost everywhere. Define function $J : B_{\nearrow} \rightarrow \mathbb{R}$,

$$J(\phi) := \int_0^1 \sqrt{\dot{\phi}(x)\rho(x, \phi(x))} dx \quad \text{and} \quad \bar{J} := \sup_{\phi \in B_{\nearrow}} J(\phi).$$

Here $\rho(x, y)$ is the density defined in (2).

Remark. By Theorems 3 and 4 in [9] it follows from Lemma 4.5 (a) and (b), that

$$\sup_{\phi \in B_{\nearrow}} J(\phi) = \sup_{\phi \in B_{\nearrow}^1} J(\phi),$$

where B_{\nearrow}^1 is defined in Theorem 1.4. Hence we use the same notation \bar{J} to denote the supremum over B_{\nearrow} .

Lemma 4.7. Under the same conditions as in Theorem 1.4, for any $\phi \in B_{\nearrow}^1$ and any $\delta, \epsilon > 0$, define the event

$$E_n := \left\{ (\pi, \tau) \in S_n \times S_n : \exists \text{ an increasing subsequence of } \left\{ \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n} \right) \right\}_{i \in [n]} \right. \\ \left. \text{which is wholly contained in the } \delta \text{ neighborhood of } \phi(\cdot) \right. \\ \left. \text{and the length of which is greater than } 2J(\phi)(1 - \epsilon)\sqrt{n} \right\}.$$

Here we say a point (x, y) is in the δ neighborhood of ϕ if $\phi(x) - \delta < y < \phi(x) + \delta$. Then

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(E_n) = 1.$$

Proof. Given $\delta, \epsilon > 0$, fix an integer K . Let $\Delta x := 1/K$. Let $x_i := i\Delta x$ and $y_i := \phi(x_i)$ for $i \in [K]$. Let $x_0 := 0$, $y_0 := 0$. Define the rectangles $R_i := [x_{i-1}, x_i] \times [y_{i-1}, y_i]$ for $i \in [K]$. Since ϕ is in C_b^1 , for any $0 < \delta' < 1$, we can choose K large enough such that

$$\max_i (y_i - y_{i-1}) < \delta, \quad e^{-\Delta x |\beta|/2} > 1 - \delta', \quad \Delta x |\beta| < \ln 2 \quad (44)$$

$$\max_i \max_{x,y \in R_i} \max \left(\frac{\rho(x,y)}{\rho(x_i,y_i)}, \frac{\rho(x_i,y_i)}{\rho(x,y)} \right) < \frac{1}{1-\delta'}, \quad (45)$$

and

$$\sum_{i=1}^K \sqrt{\rho(x_i, y_i)(y_i - y_{i-1})\Delta x} > (1 - \delta')J(\phi). \quad (46)$$

(45) follows from the uniform continuity of $\rho(x, y)$ on $[0, 1] \times [0, 1]$ and the fact that $\rho(x, y)$ is bounded away from 0, which is proved in Lemma 4.5 (a). (46) follows since

$$\begin{aligned} & \lim_{K \rightarrow \infty} \sum_{i=1}^K \sqrt{\rho(x_i, y_i)(y_i - y_{i-1})\Delta x} \\ &= \lim_{K \rightarrow \infty} \sum_{i=1}^K \sqrt{\rho(x_i, y_i) \frac{y_i - y_{i-1}}{x_i - x_{i-1}}} \Delta x \\ &= J(\phi). \end{aligned}$$

Here the last equality follows from the definition of Riemann integral, the mean value theorem and the fact that $\phi \in C_b^1$.

Next, for any $i \in [K]$, define $\rho(R_i) := \iint_{R_i} \rho(x, y) dx dy$. By (45), we have

$$\frac{\rho(R_i)}{1 - \delta'} > \rho(x_i, y_i)(y_i - y_{i-1})\Delta x.$$

Hence, for any $i \in [K]$, we have

$$\frac{l_{R_i}(\pi, \tau)}{2\sqrt{n\rho(x_i, y_i)(y_i - y_{i-1})\Delta x}} \geq \frac{l_{R_i}(\pi, \tau)\sqrt{1 - \delta'}}{2\sqrt{n\rho(R_i)}}. \quad (47)$$

By fixing the ϵ in Lemma 4.4 to be $2\delta'$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\frac{l_{R_i}(\pi, \tau)}{\sqrt{n\rho(R_i)}} > 2e^{-\Delta x|\beta|/2} - 2\delta' \right) = 1. \quad (48)$$

Moreover,

$$\mathbb{P}_n \left(\frac{l_{R_i}(\pi, \tau)}{2\sqrt{n\rho(x_i, y_i)(y_i - y_{i-1})\Delta x}} > (1 - 2\delta')\sqrt{1 - \delta'} \right) \quad (49)$$

$$\begin{aligned}
&\geq \mathbb{P}_n \left(\frac{l_{R_i}(\pi, \tau)}{2\sqrt{n\rho(R_i)}} > 1 - 2\delta' \right) \\
&\geq \mathbb{P}_n \left(\frac{l_{R_i}(\pi, \tau)}{\sqrt{n\rho(R_i)}} > 2e^{-\Delta x|\beta|/2} - 2\delta' \right).
\end{aligned}$$

The first inequality follows by (47), and the second inequality follows by (44), since

$$2e^{-\Delta x|\beta|/2} - 2\delta' > 2(1 - \delta') - 2\delta' = 2(1 - 2\delta').$$

Hence, by (48) and (49), we get

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\frac{l_{R_i}(\pi, \tau)}{2\sqrt{n\rho(x_i, y_i)(y_i - y_{i-1})\Delta x}} > (1 - 2\delta')\sqrt{1 - \delta'} \right) = 1, \quad (50)$$

for any $i \in [K]$.

Note that by concatenating the increasing subsequences of $\left\{ \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n} \right) \right\}_{i \in [n]}$ in each R_i we get a increasing subsequence in $[0, 1] \times [0, 1]$ which is wholly contained in a δ neighborhood of ϕ . Combining (46) and (50), it follows that, with probability converging to 1 as $n \rightarrow \infty$, there exists an increasing subsequence of $\left\{ \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n} \right) \right\}_{i \in [n]}$ in a δ neighborhood of ϕ whose length is at least

$$\sum_{i=1}^K 2\sqrt{n}(1 - 2\delta')\sqrt{1 - \delta'}\sqrt{\rho(x_i, y_i)(y_i - y_{i-1})\Delta x} > 2\sqrt{n}(1 - 2\delta')(1 - \delta')^{\frac{3}{2}}J(\phi).$$

The lemma follows since we can choose δ' small enough in the first place such that $(1 - 2\delta')(1 - \delta')^{\frac{3}{2}} > 1 - \epsilon$. \square

Definition 4.8. Given $K, L \in \mathbb{N}$ and multi-indices $\mathbf{b} = (b_0, b_1, \dots, b_K)$ such that $0 = b_0 \leq b_1 \leq \dots \leq b_K = KL - 1$, for any $i \in [K]$, define the rectangle $R_i := ((i - 1)\Delta x, i\Delta x] \times (b_{i-1}\Delta y, (b_i + 1)\Delta y]$, where $\Delta x := \frac{1}{K}$ and $\Delta y := \frac{1}{KL}$. Let $M_i := \sup_{(x, y) \in R_i} \rho(x, y)$ and $m_i := \inf_{(x, y) \in R_i} \rho(x, y)$. Define

$$J_{\mathbf{b}}^{K, L} := \sum_{i=1}^K \sqrt{M_i(b_i - b_{i-1} + 1)\Delta x \Delta y}.$$

Lemma 4.9.

$$\overline{\lim}_{\substack{K \rightarrow \infty \\ L \rightarrow \infty}} \max_{\mathbf{b}} J_{\mathbf{b}}^{K,L} \leq \bar{J}$$

where \bar{J} is defined in Definition 4.6, and the maximum is taken over all $\mathbf{b} = (b_0, b_1, \dots, b_K)$ such that $0 = b_0 \leq b_1 \leq \dots \leq b_K = KL - 1$.

Proof. Let M be an upper bound of $\rho(x, y)$. In the context of Definition 4.8, let $\phi_{\mathbf{b}}(x)$ be the piecewise linear function on $[0, 1]$ such that $\phi_{\mathbf{b}}(i\Delta x) = b_i\Delta y$, $i = 0, 1, \dots, K$. From the two definitions above, we have

$$\begin{aligned} J(\phi_{\mathbf{b}}) &= \int_0^1 \sqrt{\dot{\phi}_{\mathbf{b}}(x) \rho(x, \phi_{\mathbf{b}}(x))} dx \\ &= \sum_{i=1}^K \int_{(i-1)\Delta x}^{i\Delta x} \sqrt{\dot{\phi}_{\mathbf{b}}(x) \rho(x, \phi_{\mathbf{b}}(x))} dx \\ &= \sum_{i=1}^K \int_{(i-1)\Delta x}^{i\Delta x} \sqrt{\frac{(b_i - b_{i-1})\Delta y}{\Delta x} \cdot \rho(x, \phi_{\mathbf{b}}(x))} dx \\ &\geq \sum_{i=1}^K \int_{(i-1)\Delta x}^{i\Delta x} \sqrt{\frac{(b_i - b_{i-1})\Delta y}{\Delta x} \cdot m_i} dx \\ &= \sum_{i=1}^K \sqrt{m_i(b_i - b_{i-1})\Delta x \Delta y} \\ &\geq \sum_{i=1}^K \sqrt{M_i(b_i - b_{i-1})\Delta x \Delta y} - \sum_{i=1}^K \sqrt{(M_i - m_i)(b_i - b_{i-1})\Delta x \Delta y}. \end{aligned} \tag{51}$$

Here the last inequality follows since, for $a, b \geq 0$, $\sqrt{a} + \sqrt{b} \geq \sqrt{a+b}$. Moreover,

$$\begin{aligned} &\sum_{i=1}^K \sqrt{M_i(b_i - b_{i-1})\Delta x \Delta y} \\ &= J_{\mathbf{b}}^{K,L} - \sum_{i=1}^K (\sqrt{M_i(b_i - b_{i-1} + 1)\Delta x \Delta y} - \sqrt{M_i(b_i - b_{i-1})\Delta x \Delta y}) \\ &= J_{\mathbf{b}}^{K,L} - \sum_{i=1}^K \frac{M_i \Delta x \Delta y}{\sqrt{M_i(b_i - b_{i-1} + 1)\Delta x \Delta y} + \sqrt{M_i(b_i - b_{i-1})\Delta x \Delta y}} \end{aligned} \tag{52}$$

$$\begin{aligned}
&\geq J_{\mathbf{b}}^{K,L} - \sum_{i=1}^K \frac{M_i \Delta x \Delta y}{\sqrt{M_i(b_i - b_{i-1} + 1) \Delta x \Delta y}} \\
&\geq J_{\mathbf{b}}^{K,L} - \sum_{i=1}^K \frac{M_i \Delta x \Delta y}{\sqrt{M_i \Delta x \Delta y}} \\
&\geq J_{\mathbf{b}}^{K,L} - \sqrt{M} \sum_{i=1}^K \sqrt{\Delta x \Delta y} \\
&= J_{\mathbf{b}}^{K,L} - \sqrt{\frac{M}{L}}.
\end{aligned}$$

Next, define

$$\begin{aligned}
D_1(\mathbf{b}) &:= \{i \in [K] : (b_i - b_{i-1} + 1) \Delta y \leq \sqrt[3]{\Delta x}\}, \\
D_2(\mathbf{b}) &:= \{i \in [K] : (b_i - b_{i-1} + 1) \Delta y > \sqrt[3]{\Delta x}\}.
\end{aligned}$$

For $i \in D_1(\mathbf{b})$, the height of R_i is no greater than $\sqrt[3]{\Delta x}$, and for $i \in D_2(\mathbf{b})$, the height of R_i is greater than $\sqrt[3]{\Delta x}$. To bound the cardinality of $D_2(\mathbf{b})$, we have

$$\begin{aligned}
|D_2(\mathbf{b})| \sqrt[3]{\Delta x} &\leq \sum_{i \in D_2(\mathbf{b})} (b_i - b_{i-1} + 1) \Delta y \\
&\leq \sum_{i \in D_2(\mathbf{b})} (b_i - b_{i-1}) \Delta y + |D_2(\mathbf{b})| \Delta y \\
&\leq \sum_{i=1}^K (b_i - b_{i-1}) \Delta y + K \Delta y \\
&\leq 1 + \frac{1}{L} \\
&\leq 2.
\end{aligned} \tag{53}$$

Given $\epsilon > 0$, by the uniform continuity of $\rho(x, y)$ on $[0, 1] \times [0, 1]$, there exists $K_0 > 0$ such that, for any $K > K_0$ and any $i \in D_1(\mathbf{b})$, we have $M_i - m_i < \epsilon^2$. We can also choose K_0 sufficiently large such that, for any $K > K_0$,

$$2\sqrt{M}(\Delta x)^{\frac{1}{6}} < \epsilon. \tag{54}$$

Thus, for any $K > K_0$, we have

$$\sum_{i=1}^K \sqrt{(M_i - m_i)(b_i - b_{i-1}) \Delta x \Delta y} \tag{55}$$

$$\begin{aligned}
&\leq \sum_{i \in D_1(\mathbf{b})} \sqrt{\epsilon^2(b_i - b_{i-1})\Delta x \Delta y} + \sum_{i \in D_2(\mathbf{b})} \sqrt{M(b_i - b_{i-1})\Delta x \Delta y} \\
&\leq \epsilon \sum_{i=1}^K \sqrt{(b_i - b_{i-1})\Delta x \Delta y} + \sum_{i \in D_2(\mathbf{b})} \sqrt{M\Delta x} \\
&\leq \epsilon \sqrt{\sum_{i=1}^K \Delta x} \sqrt{\sum_{i=1}^K (b_i - b_{i-1})\Delta y} + 2\sqrt{M}(\Delta x)^{\frac{1}{6}} \\
&< \epsilon + \epsilon.
\end{aligned}$$

Here the second to last inequality follows by Cauchy-Schwarz inequality and (53). Let $L_0 := \lceil \frac{M}{\epsilon^2} \rceil$. By combining (51), (52) and (55), we get, for any $K > K_0$, $L > L_0$ and any \mathbf{b} ,

$$J_{\mathbf{b}}^{K,L} \leq J(\phi_{\mathbf{b}}) + \sqrt{\frac{M}{L}} \leq J(\phi_{\mathbf{b}}) + 3\epsilon \leq \bar{J} + 3\epsilon.$$

Here the last inequality follows from the fact that $\phi_{\mathbf{b}} \in B_{\nearrow}$ and Definition 4.6. \square

Definition 4.10. *In the context of Definition 4.8, we call a sequence of points (z_1, \dots, z_m) with $z_i = (x_i, y_i)$ a \mathbf{b} -increasing sequence if the following two conditions are satisfied.*

- (a) (z_1, \dots, z_m) is an increasing sequence, that is $x_i < x_{i+1}$ and $y_i < y_{i+1}$ for all $i \in [m-1]$.
- (b) Every point in the sequence is contained in some rectangle R_j with $j \in [K]$. In other words, $(j-1)\Delta x < x_i \leq j\Delta x$ implies $b_{j-1}\Delta y < y_i \leq (b_j+1)\Delta y$.

Given a collection of points $\mathbf{z} = \{z_i\}_{i \in [n]}$, let $LIS_{\mathbf{b}}(\mathbf{z})$ denote the length of the longest \mathbf{b} -increasing subsequence of \mathbf{z} . That is

$$\begin{aligned}
LIS_{\mathbf{b}}(\mathbf{z}) &:= \max\{m : \exists(i_1, i_2, \dots, i_m) \\
&\quad \text{such that } (z_{i_1}, z_{i_2}, \dots, z_{i_m}) \text{ is a } \mathbf{b}\text{-increasing sequence}\}.
\end{aligned}$$

Here we do not require $i_j < i_{j+1}$.

Lemma 4.11. *Under the same conditions as in Lemma 1.3, for any $\delta > 0$, there exist K_0, L_0 such that, for any $K > K_0$, $L > L_0$ and any $\mathbf{b} = (b_0, b_1, \dots, b_K)$ with $0 = b_0 \leq b_1 \leq \dots \leq b_K = KL - 1$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(LIS_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > 2\sqrt{n}(\bar{J} + \delta)) = 0. \quad (56)$$

Here $\mathbf{z}(\pi, \tau) := \left\{ \left(\frac{\pi(i)}{n}, \frac{\tau(i)}{n} \right) \right\}_{i \in [n]}$.

Proof. Given $\delta > 0$, by Lemma 4.9, there exist $K_1, L_1 > 0$ such that, for any $K > K_1, L > L_1$ and any $\mathbf{b} = (b_0, b_1, \dots, b_K)$ with $0 = b_0 \leq b_1 \leq \dots \leq b_K = KL - 1$, we have

$$J_{\mathbf{b}}^{K,L} < \bar{J} + \frac{\delta}{2}.$$

Then, we get

$$\mathbb{P}_n(\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > 2\sqrt{n}(\bar{J} + \delta)) \leq \mathbb{P}_n(\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > 2\sqrt{n}(J_{\mathbf{b}}^{K,L} + \delta/2)).$$

Hence, to show (56), it suffices to show that there exists K_2, L_2 such that, for any $K > K_2, L > L_2$ and any \mathbf{b} ,

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > 2\sqrt{n}(J_{\mathbf{b}}^{K,L} + \delta/2)) = 0. \quad (57)$$

Given $K, L > 0$, whose values are to be determined, and any $\mathbf{b} = (b_0, \dots, b_K)$ with $0 = b_0 \leq b_1 \leq \dots \leq b_K = KL - 1$, we inherit all the notations introduced in Definition 4.8. Let $l_{R_i}(\pi, \tau)$ denote the length of the longest increasing subsequence of $\mathbf{z}(\pi, \tau)$ wholly contained in the rectangle R_i . For any $i \in [K]$, define

$$E_i(\mathbf{b}) := \{(\pi, \tau) : l_{R_i}(\pi, \tau) \geq 2\sqrt{n}(\sqrt{M_i(b_i - b_{i-1} + 1)\Delta x \Delta y} + \delta \Delta x / 2)\}.$$

Since $\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) \leq \sum_{i=1}^K l_{R_i}(\pi, \tau)$, we get

$$\{\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > 2\sqrt{n}(J_{\mathbf{b}}^{K,L} + \delta/2)\} \subset \bigcup_{i \in [K]} E_i(\mathbf{b}).$$

Hence, to show (57), it suffices to show

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(E_i(\mathbf{b})) = 0, \quad \forall i \in [K]. \quad (58)$$

Since $e^{\Delta x |\beta|/2} - 1 = \Theta(\Delta x)$, there exists $K_2 > 0$ such that, for any $K > K_2$, we have

$$e^{\Delta x |\beta|/2} < 1 + \frac{\delta \sqrt{\Delta x}}{2\sqrt{M}} \quad \text{and} \quad \Delta x |\beta| < \ln 2. \quad (59)$$

Here $M := \sup_{0 \leq x, y \leq 1} \rho(x, y)$. Moreover, for any $i \in [K]$,

$$\mathbb{P}_n(E_i(\mathbf{b})) \quad (60)$$

$$\begin{aligned}
&\leq \mathbb{P}_n \left(l_{R_i}(\pi, \tau) \geq 2\sqrt{n} \sqrt{M_i(b_i - b_{i-1} + 1)\Delta x \Delta y} \left(1 + \frac{\delta \Delta x}{2\sqrt{M\Delta x}}\right) \right) \\
&\leq \mathbb{P}_n \left(l_{R_i}(\pi, \tau) \geq 2\sqrt{n\rho(R_i)} \left(1 + \frac{\delta \sqrt{\Delta x}}{2\sqrt{M}}\right) \right).
\end{aligned}$$

Here the first inequality follows since $(b_i - b_{i-1} + 1)\Delta y \leq 1$ and $M_i \leq M$. The second one follows since

$$M_i(b_i - b_{i-1} + 1)\Delta x \Delta y \geq \int_{R_i} \rho(x, y) dx dy = \rho(R_i).$$

Hence, combining (59), (60) and Lemma 4.4, we get, for any $K > K_2$, $L > 0$ and any \mathbf{b} ,

$$\lim_{n \rightarrow \infty} \mathbb{P}_n(E_i(\mathbf{b})) = 0, \quad \forall i \in [K].$$

Thus, (58) as well as the lemma follow. \square

Proof of Theorem 1.4. By Proposition 3.7, if $\pi \sim \mu_{n,q}$, π^{-1} has the same distribution $\mu_{n,q}$. Hence, if $(\pi, \tau) \sim \mu_{n,q} \times \mu_{n,q'}$, (π^{-1}, τ^{-1}) has the same distribution $\mu_{n,q} \times \mu_{n,q'}$. Thus, by Corollary 2.4, to prove Theorem 1.4, it suffices to show

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\left| \frac{\text{LIS}(\mathbf{z}(\pi, \tau))}{\sqrt{n}} - 2\bar{J} \right| < \epsilon \right) = 1, \quad (61)$$

for any $\epsilon > 0$. Here we use the trivial fact that $\text{LIS}(\pi, \tau) = \text{LIS}(\mathbf{z}(\pi, \tau))$. By Lemma 4.7 and the definition of \bar{J} , we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\frac{\text{LIS}(\mathbf{z}(\pi, \tau))}{\sqrt{n}} > 2\bar{J} - \epsilon \right) = 1. \quad (62)$$

To show the upper bound in (61), note that, for any $K, L > 0$ and any increasing sequence of points $\{(x_j, y_j)\}_{j \in [n]}$ with $0 < x_j, y_j \leq 1$, there exists a choice of $\mathbf{b} = (b_0, b_1, \dots, b_K)$ such that $\{(x_j, y_j)\}_{j \in [n]}$ is a \mathbf{b} -increasing sequence. Specifically, we can define \mathbf{b} as follows. Let $\Delta x := \frac{1}{K}$, $\Delta y := \frac{1}{KL}$.

- Define $b_0 := 0$, $b_K := KL - 1$.
- For $i \in [K - 1]$, define $b_i := \lfloor \max \{y_j : (i - 1)\Delta x < x_j \leq i\Delta x\} \cdot KL \rfloor$.

It can be easily verified that with \mathbf{b} thus defined, every point (x_j, y_j) is in some rectangle R_i , where R_i is defined in Definition 4.8. Hence, we get

$$\mathbb{P}_n \left(\frac{\text{LIS}(\mathbf{z}(\pi, \tau))}{\sqrt{n}} > 2\bar{J} + \epsilon \right) \quad (63)$$

$$\begin{aligned}
&= \mathbb{P}_n \left(\max_{\mathbf{b}} (\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau))) > \sqrt{n}(2\bar{J} + \epsilon) \right) \\
&\leq \sum_{\mathbf{b}} \mathbb{P}_n (\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > \sqrt{n}(2\bar{J} + \epsilon)).
\end{aligned}$$

Here, the maximum and summation are taken over all possible \mathbf{b} with $0 = b_0 \leq b_1 \leq \dots \leq b_K = KL - 1$. By Lemma 4.11, we can choose K, L sufficiently large such that, for any \mathbf{b} ,

$$\lim_{n \rightarrow \infty} \mathbb{P}_n (\text{LIS}_{\mathbf{b}}(\mathbf{z}(\pi, \tau)) > \sqrt{n}(2\bar{J} + \epsilon)) = 0.$$

Hence, by (63) and the fact that the number of different choices of \mathbf{b} is bounded above by $(KL)^K$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\frac{\text{LIS}(\mathbf{z}(\pi, \tau))}{\sqrt{n}} > 2\bar{J} + \epsilon \right) = 0. \quad (64)$$

And (61) follows from (62) and (64) □

The following lemma let us solve for the supremum \bar{J} when the underlying density $\rho(x, y)$ satisfies $\rho\left(\frac{x+y}{2}, \frac{x+y}{2}\right) \geq \rho(x, y)$.

Lemma 4.12. *Given a density $\rho(x, y)$ on $[0, 1] \times [0, 1]$ such that $\rho(x, y)$ is C_b^1 and $c < \rho(x, y) < C$ for some $C, c > 0$, if $\rho(x, y) \leq \rho\left(\frac{x+y}{2}, \frac{x+y}{2}\right)$ for any $0 \leq x, y \leq 1$, then we have*

$$\bar{J} = \int_0^1 \sqrt{\rho(x, x)} dx,$$

i. e. the supremum of $J(\phi)$ on B_{\nearrow} is attained for $\phi(x) = x$.

Proof. By the remark following Definition 4.6, it suffices to show that, for any $\phi \in B_{\nearrow}^1$, we have

$$J(\phi) \leq \int_0^1 \sqrt{\rho(x, x)} dx. \quad (65)$$

Define $g_\phi(x) := x + \phi(x)$. Since $\dot{\phi}(x) \geq 0$, we have $\dot{g}_\phi(x) \geq 1$. Next, we reparameterize $\phi(x)$ as follows,

$$t := \frac{g_\phi(x)}{2} = \frac{x + \phi(x)}{2}. \quad (66)$$

Thus, we have $x = g_\phi^{-1}(2t)$ and $\phi(x) = 2t - x = 2t - g_\phi^{-1}(2t)$ where $t \in [0, 1]$. Moreover, since $g_\phi(x)$ is strictly increasing, x is strictly increasing as a function of t . Hence we have

$$\rho(x, \phi(x)) = \rho(g_\phi^{-1}(2t), 2t - g_\phi^{-1}(2t)) \leq \rho(t, t). \quad (67)$$

Here the last inequality follows since $\rho(x, y) \leq \rho\left(\frac{x+y}{2}, \frac{x+y}{2}\right)$. Next, by taking derivative with respect to t on both sides of (66), we have

$$1 = \frac{1}{2} \left(\frac{dx}{dt} + \dot{\phi}(x) \frac{dx}{dt} \right). \quad (68)$$

By multiplying $2 \frac{dx}{dt}$ on both sides of (68), we get

$$\dot{\phi}(x) \left(\frac{dx}{dt} \right)^2 = 2 \frac{dx}{dt} - \left(\frac{dx}{dt} \right)^2 \leq 1. \quad (69)$$

Hence, by (67) and (69), we have

$$\begin{aligned} J(\phi) &= \int_0^1 \sqrt{\dot{\phi}(x) \rho(x, \phi(x))} dx \\ &\leq \int_0^1 \sqrt{\dot{\phi}(x) \rho(t, t)} \cdot \frac{dx}{dt} dt \\ &= \int_0^1 \sqrt{\rho(t, t) \dot{\phi}(x) \left(\frac{dx}{dt} \right)^2} dt \\ &\leq \int_0^1 \sqrt{\rho(t, t)} dt. \end{aligned}$$

Therefore, \bar{J} is attained for $\phi(x) = x$. \square

Proof of Corollary 1.5. Note that in the special case where $\beta = \gamma$, the density $\rho(x, y)$ in (2) is given by

$$\rho(x, y) := \int_0^1 u(x, t, \beta) \cdot u(t, y, \beta) dt. \quad (70)$$

In this case, we will show that $\rho(x, y) \leq \rho\left(\frac{x+y}{2}, \frac{x+y}{2}\right)$ for any $0 \leq x, y \leq 1$. Hence, by Lemma 4.5 and Lemma 4.12, \bar{J} defined in Theorem 1.4 is attained when $\phi(x) = x$. In fact, by direct calculation, it can be shown that

$$u(x, t, \beta) \cdot u(t, y, \beta) \leq u\left(\frac{x+y}{2}, t, \beta\right) \cdot u\left(t, \frac{x+y}{2}, \beta\right), \quad (71)$$

for any $0 \leq x, y, t \leq 1$.

By the definition of $u(x, y, \beta)$, we have

$$\begin{aligned}
& u(x, t, \beta) \cdot u(t, y, \beta) \\
&= \frac{(\beta/2) \sinh(\beta/2)}{(e^{\beta/4} \cosh(\beta[x-t]/2) - e^{-\beta/4} \cosh(\beta[x+t-1]/2))^2} \\
&\quad \times \frac{(\beta/2) \sinh(\beta/2)}{(e^{\beta/4} \cosh(\beta[t-y]/2) - e^{-\beta/4} \cosh(\beta[t+y-1]/2))^2} \\
&= \frac{\beta(e^\beta - 1)}{(2e^{\beta/2} \cosh(\beta[x-t]/2) - 2 \cosh(\beta[x+t-1]/2))^2} \\
&\quad \times \frac{\beta(e^\beta - 1)}{(2e^{\beta/2} \cosh(\beta[t-y]/2) - 2 \cosh(\beta[t+y-1]/2))^2}.
\end{aligned} \tag{72}$$

Considering the term inside the square of the denominator, by using the hyperbolic trigonometric identities,

$$\begin{aligned}
\cosh(x) \cosh(y) &= (\cosh(x+y) + \cosh(x-y))/2, \\
\cosh(x+y) &= \cosh(x) \cosh(y) + \sinh(x) \sinh(y), \\
\cosh(x-y) &= \cosh(x) \cosh(y) - \sinh(x) \sinh(y),
\end{aligned}$$

we get

$$\begin{aligned}
& (2e^{\beta/2} \cosh(\beta[x-t]/2) - 2 \cosh(\beta[x+t-1]/2)) \\
&\quad \times (2e^{\beta/2} \cosh(\beta[t-y]/2) - 2 \cosh(\beta[t+y-1]/2)) \\
&= 2e^\beta (\cosh(\beta[x-y]/2) + \cosh(\beta[x+y-2t]/2)) \\
&\quad - 2e^{\beta/2} (\cosh(\beta[x+y-1]/2) + \cosh(\beta[x-y-2t+1]/2)) \\
&\quad - 2e^{\beta/2} (\cosh(\beta[x-y+2t-1]/2) + \cosh(\beta[x+y-1]/2)) \\
&\quad + 2 (\cosh(\beta[x+y+2t-2]/2) + \cosh(\beta[x-y]/2)) \\
&= S_t^- + S_t^+.
\end{aligned} \tag{73}$$

Here S_t^- denotes the sum of those terms in the above equation containing the term $x-y$. And S_t^+ denotes the sum of those which contain the term $x+y$. After further simplification using the identities above, we have

$$S_t^- = 2 \cosh(\beta[x-y]/2) (e^\beta - 2e^{\beta/2} \cosh(\beta[2t-1]/2) + 1). \tag{74}$$

It is easily seen that the minimum of $e^\beta - 2e^{\beta/2} \cosh(\beta[2t - 1]/2) + 1$ for $0 \leq t \leq 1$ is attained when $t = 0, 1$, and the minimum is 0. Hence, for any $t \in [0, 1]$, S_t^- is minimized when $x = y$. Thus to prove (71), it suffices to show that $S_t^+ \geq 0$, since $S_t^- + S_t^+$ is the term inside the square of the denominator of (72). After simplification, we have

$$\begin{aligned} S_t^+ = & 2e^\beta \left(\cosh(\beta[x + y - 1]/2) \cosh(\beta[2t - 1]/2) \right. \\ & \left. - \sinh(\beta[x + y - 1]/2) \sinh(\beta[2t - 1]/2) \right) \\ & - 4e^{\beta/2} \cosh(\beta[x + y - 1]/2) \\ & + 2 \left(\cosh(\beta[x + y - 1]/2) \cosh(\beta[2t - 1]/2) \right. \\ & \left. + \sinh(\beta[x + y - 1]/2) \sinh(\beta[2t - 1]/2) \right). \end{aligned} \quad (75)$$

Next, we make change of variables. Define $r := e^{\beta(x+y-1)/2}$, $s := e^{\beta(2t-1)/2}$. Then, from (75), we have

$$\begin{aligned} S_t^+ = & \frac{e^\beta}{2} \left(\left(r + \frac{1}{r} \right) \left(s + \frac{1}{s} \right) - \left(r - \frac{1}{r} \right) \left(s - \frac{1}{s} \right) \right) - 2e^{\beta/2} \left(r + \frac{1}{r} \right) \\ & + \frac{1}{2} \left(\left(r + \frac{1}{r} \right) \left(s + \frac{1}{s} \right) + \left(r - \frac{1}{r} \right) \left(s - \frac{1}{s} \right) \right) \\ = & e^\beta \left(\frac{r}{s} + \frac{s}{r} \right) - 2e^{\beta/2} \left(r + \frac{1}{r} \right) + \left(rs + \frac{1}{rs} \right) \\ = & \left(\frac{e^\beta r}{s} + rs - 2e^{\beta/2} r \right) + \left(\frac{e^\beta s}{r} + \frac{1}{rs} - \frac{2e^{\beta/2}}{r} \right) \\ \geq & 0. \end{aligned} \quad (76)$$

Here the last inequality follows since $x + y \geq 2\sqrt{xy}$ for any $x, y \geq 0$. We complete the proof of Corollary 1.5 by showing:

$$\int_0^1 u(x, t, \beta) \cdot u(t, x, \beta) dt = \frac{\beta \left(\cosh(\beta/2) + 2 \cosh(\beta[2x - 1]/2) \right)}{6 \sinh(\beta/2)}, \quad (77)$$

for $0 \leq x \leq 1$.

By the same change of variables as above, since $y = x$, let $r := e^{\beta(2x-1)/2}$, $s := e^{\beta(2t-1)/2}$. Then, we have

$$\frac{dt}{ds} = \frac{1}{\frac{ds}{dt}} = \frac{1}{s\beta}. \quad (78)$$

By (74), we have,

$$S_t^- = 2 \left(e^\beta - e^{\beta/2} \left(s + \frac{1}{s} \right) + 1 \right). \quad (79)$$

Then, by (76) and (79), it can be easily verified that

$$rs (S_t^+ + S_t^-) = (e^{\beta/2}(r+s) - (rs+1))^2. \quad (80)$$

Hence, we have

$$\begin{aligned} & \int_0^1 u(x, t, \beta) \cdot u(t, x, \beta) dt \\ &= \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{\beta^2 (e^\beta - 1)^2}{(S_t^+ + S_t^-)^2} \frac{1}{s\beta} ds \\ &= \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{\beta (e^\beta - 1)^2 r^2 s}{(rs (S_t^+ + S_t^-))^2} ds \\ &= \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{\beta (e^\beta - 1)^2 r^2 s}{(e^{\beta/2}(r+s) - (rs+1))^4} ds \\ &= \beta (e^\beta - 1)^2 r^2 \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{s}{((e^{\beta/2} - r)s + e^{\beta/2}r - 1)^4} ds \\ &= \beta (e^\beta - 1)^2 e^{\beta(2x-1)} \int_{e^{-\beta/2}}^{e^{\beta/2}} \frac{s}{(e^{\beta/2}(1 - e^{\beta(x-1)})s + e^{\beta x} - 1)^4} ds. \end{aligned} \quad (81)$$

Here the first equality follows from (72), (73), (78) and change of variables. The third equality follows from (80). Then we make another change of variable by defining

$$w := \frac{e^{\beta/2}(1 - e^{\beta(x-1)})s + e^{\beta x} - 1}{e^\beta - 1},$$

from which we have

$$\frac{ds}{dw} = \frac{e^\beta - 1}{e^{\beta/2}(1 - e^{\beta(x-1)})}, \quad \text{and} \quad w = \begin{cases} 1 & \text{when } s = e^{\beta/2}, \\ e^{\beta(x-1)} & \text{when } s = e^{-\beta/2}. \end{cases}$$

Hence, by (81), we have

$$\int_0^1 u(x, t, \beta) \cdot u(t, x, \beta) dt$$

$$\begin{aligned}
&= \frac{\beta e^{2\beta(x-1)}}{(e^\beta - 1)(1 - e^{\beta(x-1)})^2} \int_{e^{\beta(x-1)}}^1 \frac{(e^\beta - 1)w - e^{\beta x} + 1}{w^4} dw \\
&= \frac{\beta e^{2\beta(x-1)}}{(e^\beta - 1)(1 - e^{\beta(x-1)})^2} \left(\frac{1 - e^\beta}{2w^2} + \frac{e^{\beta x} - 1}{3w^3} \right) \Big|_{e^{\beta(x-1)}}^1 \\
&= \frac{\beta (1 + e^\beta + 2e^{\beta x} + 2e^{-\beta(x-1)})}{6(e^\beta - 1)} \\
&= \frac{\beta (\cosh(\beta/2) + 2 \cosh(\beta[2x - 1]/2))}{6 \sinh(\beta/2)}.
\end{aligned}$$

□

5 Proof of Lemma 4.1 and Lemma 4.2

We first introduce two corollaries of these two lemmas. Recall that, for any positive integer n and $m \in [n]$, we define

$$Q(n, m) := \{(b_1, b_2, \dots, b_m) : b_i \in [n] \text{ and } b_i < b_{i+1} \text{ for all } i\}.$$

By choosing $k_n = n$ in these two lemmas and the fact that $Q(n, n) = \{(1, 2, \dots, n)\}$, we can recover Corollary 4.3 of [19]:

Corollary 5.1. *Suppose that $\{q_n\}_{n=1}^\infty$ is a sequence such that $q_n \geq 1$ and $\liminf_{n \rightarrow \infty} n(1 - q_n) = \beta$, with $\beta \in \mathbb{R}$. We have*

$$\lim_{n \rightarrow \infty} \mu_{n, q_n} \left(\pi \in S_n : \frac{LIS(\pi)}{\sqrt{n}} \notin (2e^{\frac{\beta}{2}} - \epsilon, 2 + \epsilon) \right) = 0,$$

for any $\epsilon > 0$.

Corollary 5.2. *Suppose that $\{q_n\}_{n=1}^\infty$ is a sequence such that $0 < q_n \leq 1$ and $\limsup_{n \rightarrow \infty} n(1 - q_n) = \beta < \ln 2$. We have*

$$\lim_{n \rightarrow \infty} \mu_{n, q_n} \left(\pi \in S_n : \frac{LIS(\pi)}{\sqrt{n}} \notin (2 - \epsilon, 2e^{\frac{\beta}{2}} + \epsilon) \right) = 0,$$

for any $\epsilon > 0$.

To prove these two lemmas, we use the same techniques developed in the proof of Corollary 4.3 in [19], in which they constructed a coupling of two

point processes. A point process is a random, locally finite, nonnegative integer valued measure. Let \mathcal{X}_k denote the set of all Borel measures ξ on \mathbb{R}^k such that $\xi(A) \in \{0, 1, 2, \dots\}$ for any bounded Borel set A in \mathbb{R}^k . Then, a point process on \mathbb{R}^k is a random variable which takes value in \mathcal{X}_k . Suppose μ, ν are two measures on \mathbb{R}^k . We say $\mu \leq \nu$ if $\mu(A) \leq \nu(A)$ for any $A \in \mathcal{B}(\mathbb{R}^k)$.

Lemma 5.3. *Suppose $\hat{\alpha}$ and α are two measures on $[0, 1]$ with density $f(x)$, $g(x)$ respectively. If, for any $x \in [0, 1]$, $f(x) \geq p \cdot g(x)$ for some $1 > p > 0$, then there exist random variables X, Y and B_p such that the following hold.*

- X is $\hat{\alpha}$ -distributed, Y is α -distributed and B_p is Bernoulli distributed with $\mathbb{P}(B_p = 1) = p$.
- B_p and Y are independent.
- Define two point processes η, ξ on $[0, 1]$ as follows,

$$\xi(A) := \mathbb{1}_A(X) \quad \text{and} \quad \eta(A) := B_p \cdot \mathbb{1}_A(Y), \quad \forall A \in \mathcal{B}([0, 1]).$$

Then, we have $\eta \leq \xi$ almost surely.

Proof. Let Y, Y' and B_p be independent random variables defined on the same probability space such that Y is α -distributed, B_p is Bernoulli distributed with $\mathbb{P}(B_p = 1) = p$ and the density of the distribution of Y' is $\frac{f(x) - p \cdot g(x)}{1 - p}$. Define $X = B_p Y + (1 - B_p) Y'$. It can be easily verified that X thus defined is $\hat{\alpha}$ -distributed. Because,

$$\mathbb{P}(X \in A) = p \int_A g(x) dx + (1 - p) \int_A \frac{f(x) - p \cdot g(x)}{1 - p} dx = \int_A f(x) dx,$$

for any $A \in \mathcal{B}([0, 1])$. Finally, the two point processes ξ and η thus defined satisfy $\eta \leq \xi$, since for any $A \in \mathcal{B}([0, 1])$, when $B_p = 1$, we have $\xi(A) = \eta(A)$, and, when $B_p = 0$, we have $\eta(A) = 0$. \square

Lemma 5.4. *Suppose $\hat{\alpha}$ and α are two measures on $[0, 1]$ with density $f(x)$, $g(x)$ respectively. If, for any $x \in [0, 1]$, $(1 - \theta_1)g(x) \leq f(x) \leq (1 + \theta_2)g(x)$ for some $\theta_1, \theta_2 \geq 0$ with $\theta_1 + \theta_2 < 1$, then there exist random variables X, Y, Z and B_θ such that the following hold.*

- X is $\hat{\alpha}$ -distributed, Y and Z are α -distributed and B_θ is Bernoulli distributed with $\mathbb{P}(B_\theta = 1) = \theta$, where $\theta = \theta_1 + \theta_2$.

- B_θ, Y and Z are independent.
- Define two point processes ξ, ζ on $[0, 1]$ as follows,

$$\xi(A) := \mathbb{1}_A(X) \quad \text{and} \quad \zeta(A) := \mathbb{1}_A(Y) + B_\theta \cdot \mathbb{1}_A(Z), \quad \forall A \in \mathcal{B}([0, 1]).$$

Then, we have $\xi \leq \zeta$ almost surely.

Proof. Let Y, Z and B_θ be independent random variables defined on the same probability space such that Y, Z is α -distributed, B_θ is Bernoulli distributed with $\mathbb{P}(B_\theta = 1) = \theta$. We define a new random variable X as follows. Conditioned on $Y = y$ and $Z = z$,

- if $B_\theta = 0$, define $X = y$
- if $B_\theta = 1$, we flip a coin W with probability of head being $\frac{f(z) - (1 - \theta_1)g(z)}{\theta \cdot g(z)}$. If the result is head, define $X = z$. Else, define $X = y$.

Note that, without loss of generality, here we may assume $g(z) > 0$, since $\mathbb{P}(g(Z) = 0) = 0$. It is straight forward that the two point processes ξ and ζ thus defined satisfy $\xi \leq \zeta$ a.s.. We complete the proof by verifying that X thus defined has distribution $f(x)$.

For any $A \in \mathcal{B}([0, 1])$, the event $\{X \in A\}$ can be partitioned into three parts: $\{B_\theta = 0, Y \in A\}$, $\{B_\theta = 1, W \text{ is head}, Z \in A\}$ and $\{B_\theta = 1, W \text{ is tail}, Y \in A\}$. We have

$$\mathbb{P}(\{B_\theta = 0, Y \in A\}) = (1 - \theta) \int_A g(x) dx = (1 - \theta)\alpha(A),$$

$$\begin{aligned} & \mathbb{P}(\{B_\theta = 1, W \text{ is head}, Z \in A\}) \\ &= \theta \int_A \frac{f(z) - (1 - \theta_1)g(z)}{\theta \cdot g(z)} g(z) dz \\ &= \int_A f(z) dz - (1 - \theta_1)\alpha(A), \end{aligned}$$

$$\begin{aligned} & \mathbb{P}(\{B_\theta = 1, W \text{ is tail}, Y \in A\}) \\ &= \theta \int_A g(y) dy \int_0^1 \left(1 - \frac{f(z) - (1 - \theta_1)g(z)}{\theta \cdot g(z)}\right) g(z) dz \\ &= \alpha(A) \int_0^1 (1 + \theta_2)g(z) - f(z) dz \\ &= \alpha(A) \theta_2. \end{aligned}$$

Here we evaluate the last two probabilities by conditioning on the value of Z . Summing up the three probabilities, we get

$$\mathbb{P}(\{X \in A\}) = \int_A f(z) dz.$$

□

Next, we define a triangular array of random variables in $[0, 1]$.

Definition 5.5. Suppose that $\{q_n\}_{n=1}^\infty$ is a sequence such that $q_n > 0$. For any $n \in \mathbb{N}$, we define the random vector $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ as follows. Let $\{Y_i\}_{i=1}^n$ be i. i. d. uniform random variables on $[0, 1]$. Let $\{Y_{(i)}\}_{i=1}^n$ be the order statistics of $\{Y_i\}_{i=1}^n$. Independently, let π be a μ_{n, q_n} -distributed random variable on S_n . We define $Y_i^{(n)} := Y_{(\pi(i))}$ for all $i \in [n]$.

In the remainder of this paper, we use $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ specifically to denote the random vector defined as above. Also, we define the function Φ which maps vectors in \mathbb{R}^n or n points in \mathbb{R}^2 to the induced permutation in S_n .

Definition 5.6. Suppose $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is a vector in \mathbb{R}^n such that all its entries are different. Let $\Phi(\mathbf{x})$ denote the permutation in S_n such that, for any $i \in [n]$, $\Phi(\mathbf{x})(i) = j$ if x_i is the j -th smallest entry in \mathbf{x} . Similarly, suppose $\mathbf{z} = \{(x_i, y_i)\}_{i=1}^n$ are n points in \mathbb{R}^2 such that they share no x coordinate nor any y coordinate. Let $\Phi(\mathbf{z})$ denote the permutation in S_n such that, for any $i \in [n]$, $\Phi(\mathbf{z})(i) = j$ if there exists $k \in [n]$, such that x_k is the i -th smallest term in $\{x_i\}_{i=1}^n$ and y_k is the j -th smallest term in $\{y_i\}_{i=1}^n$.

Remark. From the above definitions, it can be easily seen that

- (a) For any $x_1 < x_2 < \dots < x_n$ and $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$, we have $\Phi(\mathbf{y}) = \Phi(\{(x_i, y_i)\}_{i=1}^n)$.
- (b) For any $\mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ and $\mathbf{b} = (b_1, b_2, \dots, b_m) \in Q(n, m)$, we have $\Phi(\mathbf{y})_{\mathbf{b}} = \Phi((y_{b_1}, y_{b_2}, \dots, y_{b_m}))$.
- (c) $\Phi((Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}))$ is μ_{n, q_n} -distributed.

It is not hard to show that the density function of $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ is the following (in the sense of a. s.),

$$f_n(\mathbf{y}) = \mu_{n,q_n}(\Phi(\mathbf{y})) \cdot n! \quad \text{for all } \mathbf{y} \in [0, 1]^n \setminus \text{Diagonal}.$$

Here the set *Diagonal* consists of all those vectors which contain (at least two) identical entries. Since $\{Y_i^{(n)}\}_{i=1}^n = \{Y_i\}_{i=1}^n$ are n i.i.d. uniform samples from $[0, 1]$, we have $\mathbb{P}((Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}) \in \text{Diagonal}) = 0$. Intuitively, for any $0 \leq y_1 < \dots < y_n \leq 1$, there are $n!$ ways to choose the vector (Y_1, \dots, Y_n) such that $\{Y_i\}_{i=1}^n = \{y_i\}_{i=1}^n$. Conditioned on $\{Y_i\}_{i=1}^n = \{y_i\}_{i=1}^n$, the probability of $(Y_1^{(n)}, \dots, Y_n^{(n)}) = (y_{\pi(1)}, \dots, y_{\pi(n)})$ is $\mu_{n,q_n}(\pi)$.

Lemma 5.7. *Given $i \in [n]$ and a vector $(y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n) \in [0, 1]^{n-1} \setminus \text{Diagonal}$, let $\hat{\alpha}$ denote the distribution of $Y_i^{(n)}$ conditioned on the event $\{Y_j^{(n)} = y_j \text{ for all } j \in [n] \setminus \{i\}\}$. Then $\hat{\alpha}$ has density $f(y)$ on $[0, 1]$ such that, excluding a set G of measure zero, for any $y, y' \in [0, 1] \setminus G$, we have*

$$f(y) \geq \min \left(q_n^n, \frac{1}{q_n^n} \right), \quad f(y) - f(y') \leq \max \left(q_n^n, \frac{1}{q_n^n} \right) - 1.$$

Proof. Since $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ has density $f_n(\mathbf{y}) = \mu_{n,q_n}(\Phi(\mathbf{y})) \cdot n!$ on $[0, 1]^n \setminus \text{Diagonal}$, the density $f(y)$ of $\hat{\alpha}$ is given by

$$f(y) = \frac{\mu_{n,q_n}(\Phi((y_1, y_2, \dots, y_{i-1}, y, y_{i+1}, \dots, y_n)))}{\int_0^1 \mu_{n,q_n}(\Phi((y_1, y_2, \dots, y_{i-1}, t, y_{i+1}, \dots, y_n))) dt},$$

for any $y \in [0, 1] \setminus \{y_1, y_2, \dots, y_{i-1}, y_{i+1}, \dots, y_n\}$.

It can be seen from the definition that $f(y)$ is a simple function which takes at most n different values. Let M and m denote the maximum and minimum of $f(y)$ respectively. Then we have $M \geq 1$ and $0 < m \leq 1$. Moreover, for any $y, y' \in [0, 1]$, let $\mathbf{y} := (y_1, y_2, \dots, y_{i-1}, y, y_{i+1}, \dots, y_n)$ and $\mathbf{y}' := (y_1, y_2, \dots, y_{i-1}, y', y_{i+1}, \dots, y_n)$. We have

$$|l(\Phi(\mathbf{y})) - l(\Phi(\mathbf{y}'))| \leq n - 1.$$

That is, if \mathbf{y} and \mathbf{y}' differ at one entry, the number of inversions of the induced permutations differ at most by $n - 1$. Hence, assuming $q_n \geq 1$, for any $y, y' \in [0, 1]$, we have

$$\frac{1}{q_n^{n-1}} \leq \frac{f(y)}{f(y')} \leq q_n^{n-1}.$$

Choose y' such that $f(y') = M$, we have $f(y) \geq M/q_n^{n-1} \geq 1/q_n^n$. For the second part, we choose y, y' such that $f(y) = M$ and $f(y') = m$. Then we have $M/m - 1 \leq q_n^{n-1} - 1 \leq q_n^n - 1$. Thus, $M - m \leq q_n^n - 1$, since $0 < m \leq 1$. On the other hand, assuming $0 < q_n < 1$, by the similar argument it follows that for any $y, y' \in [0, 1]$, we have

$$q_n^{n-1} \leq \frac{f(y)}{f(y')} \leq \frac{1}{q_n^{n-1}}.$$

Choose y' such that $f(y') = M$, we have $f(y) \geq Mq_n^{n-1} \geq q_n^n$. For the second part, we choose y, y' such that $f(y) = M$ and $f(y') = m$. Then we have $M/m - 1 \leq 1/q_n^{n-1} - 1 \leq 1/q_n^n - 1$. Thus, $M - m \leq 1/q_n^n - 1$, since $0 < m \leq 1$.

Combining the two cases above, the lemma follows. \square

Lemma 5.8. *Given $n \in \mathbb{N}$ and $q_n \geq 1$, for any $m \leq n$ and any $\mathbf{b} = (b_1, b_2, \dots, b_m) \in Q(n, m)$, there exists a random vector $(V_1, V_2, \dots, V_n) \in [0, 1]^n$ and $2m$ independent random variables $\{U_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$ such that (V_1, V_2, \dots, V_n) has the same distribution as $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$, each U_i is uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \frac{1}{q_n}$. Moreover, if we define two point processes as follows,*

$$\xi_{\mathbf{b}}^{(n)}(A) := \sum_{i=1}^m \mathbf{1}_A((i, V_{b_i})), \quad \eta_m(A) := \sum_{i=1}^m B_i \cdot \mathbf{1}_A((i, U_i)), \quad \forall A \in \mathcal{B}(\mathbb{N} \times [0, 1]),$$

we have $\eta_m \leq \xi_{\mathbf{b}}^{(n)}$ almost surely.

Proof. Given n, m and $\mathbf{b} \in Q(n, m)$, define $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ as in Definition 5.5 and, independently, define $2m$ independent random variables $\{U_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$ such that each U_i is uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \frac{1}{q_n}$. We define the random vector (V_1, V_2, \dots, V_n) as follows,

- Sample the random vector $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$, say, we get $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)}) = (y_1, y_2, \dots, y_n)$.
- For $j \in [n] \setminus \{b_i\}_{i=1}^m$, let $V_j := y_j$.

- For each $i \in [m]$, we resample $Y_{b_i}^{(n)}$ one by one, conditioned on the current value of other $Y_j^{(n)}$. Let y'_{b_i} denote the new value of $Y_{b_i}^{(n)}$ after the resampling and define $V_{b_i} := y'_{b_i}$. Specifically, for each $i \in [m]$, we sample a value y'_{b_i} according to the distribution of $Y_{b_i}^{(n)}$, conditioned on the event

$$\left\{ Y_{b_j}^{(n)} = y'_{b_j} \text{ for } \forall j < i \text{ and } Y_k^{(n)} = y_k \text{ for } \forall k \in [n] \setminus \{b_j\}_{j \in [i]} \right\}.$$

- In each resampling step, say, resampling $Y_{b_i}^{(n)}$, let $\hat{\alpha}$ denote the above conditional distribution of $Y_{b_i}^{(n)}$. By Lemma 5.7, we know that $\hat{\alpha}$ has density $f(y)$ with $f(y) \geq 1/q_n^n$ almost surely. Hence, we can couple this resampling procedure with variables U_i and B_i in the same fashion as in the proof of Lemma 5.3, with α in that lemma being the uniform measure on $[0, 1]$. Thus we have $\mathbb{1}_A((i, V_{b_i})) \geq B_i \cdot \mathbb{1}_A((i, U_i))$ a.s. for any $A \in \mathcal{B}(\mathbb{N} \times [0, 1])$.

It can be easily seen from the above procedure that (V_1, V_2, \dots, V_n) thus defined has the same distribution as $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$, and

$$\eta_m(A) = \sum_{i=1}^m B_i \cdot \mathbb{1}_A((i, U_i)) \leq \sum_{i=1}^m \mathbb{1}_A((i, V_{b_i})) = \xi_{\mathbf{b}}^{(n)}(A) \text{ a.s.}$$

for any $A \in \mathcal{B}(\mathbb{N} \times [0, 1])$.

□

Lemma 5.9. *Given $n \in \mathbb{N}$ and $0 < q_n \leq 1$ such that $q_n^n > \frac{1}{2}$, for any $m \leq n$ and any $\mathbf{b} = (b_1, b_2, \dots, b_m) \in Q(n, m)$, there exists a random vector $(V_1, V_2, \dots, V_n) \in [0, 1]^n$ and $3m$ independent random variables $\{U_i\}_{i=1}^m \cup \{U'_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$ such that (V_1, V_2, \dots, V_n) has the same distribution as $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$, each U_i, U'_i are uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = \frac{1}{q_n^n} - 1$. Moreover, if we define two point processes as follows,*

$$\begin{aligned} \xi_{\mathbf{b}}^{(n)}(A) &:= \sum_{i=1}^m \mathbb{1}_A((i, V_{b_i})), & \forall A \in \mathcal{B}(\mathbb{N} \times [0, 1]) \\ \zeta_m(A) &:= \sum_{i=1}^m \mathbb{1}_A((i, U'_i)) + B_i \cdot \mathbb{1}_A((i, U_i)), & \forall A \in \mathcal{B}(\mathbb{N} \times [0, 1]) \end{aligned}$$

we have $\xi_{\mathbf{b}}^{(n)} \leq \zeta_m$ almost surely.

Proof. The proof of this lemma is similar to the proof of Lemma 5.8. Given n, m and $\mathbf{b} \in Q(n, m)$, define $(Y_1^{(n)}, Y_2^{(n)}, \dots, Y_n^{(n)})$ as in Definition 5.5 and, independently, define $3m$ independent random variables $\{U_i\}_{i=1}^m \cup \{U'_i\}_{i=1}^m \cup \{B_i\}_{i=1}^m$ such that each U_i, U'_i are uniformly distributed on $[0, 1]$ and each B_i is a Bernoulli random variable with $\mathbb{P}(B_i = 1) = 1/q_n^n - 1$. Then we define the random vector (V_1, V_2, \dots, V_n) by the same steps as in the proof of Lemma 5.8, except that, in each resampling step, we couple the resampling of $Y_{b_i}^{(n)}$ with the variables U_i, U'_i and B_i in the same fashion as in the proof of Lemma 5.4, with α in that lemma being the uniform measure on $[0, 1]$. Note that the second inequality in Lemma 5.7 ensures that the conditions in Lemma 5.4 are met. Specifically, in each resampling step, let $f(y)$ denote the density of the conditional distribution of $Y_{b_i}^{(n)}$. Let M, m be the maximum and minimum of $f(y)$ respectively. Define $\theta_1 := 1 - m$ and $\theta_2 := M - 1$. Hence, $1 - \theta_1 \leq f(y) \leq 1 + \theta_2$ almost surely and $\theta_1 + \theta_2 = M - m \leq 1/q_n^n - 1 < 1$. \square

Recall that \mathcal{X}_2 denotes the set of all Borel measures ξ on \mathbb{R}^2 such that $\xi(A) \in \{0, 1, 2, \dots\}$ for any bounded Borel set A in \mathbb{R}^2 .

Definition 5.10. For any $\xi \in \mathcal{X}_2$, we define the length of the longest increasing subsequence of ξ as follows,

$$\text{LIS}(\xi) := \max\{k : \exists (x_1, y_1), (x_2, y_2), \dots, (x_k, y_k) \in \mathbb{R}^2 \text{ such that} \\ \xi(\{(x_i, y_i)\}) \geq 1, \forall i \in [k] \text{ and } (x_i - x_j)(y_i - y_j) > 0, \forall i \neq j\}.$$

It is easily seen that the function $\text{LIS}()$ is non-decreasing on \mathcal{X}_2 in the sense that, if $\xi, \zeta \in \mathcal{X}_2$ with $\xi \leq \zeta$, we have $\text{LIS}(\xi) \leq \text{LIS}(\zeta)$. Moreover, for any n points $\{(x_i, y_i)\}_{i=1}^n$ in \mathbb{R}^2 such that $x_i \neq x_j$ and $y_i \neq y_j$ for all $i \neq j$, define the integer-valued measure ξ as follows,

$$\xi(A) := \sum_{i=1}^n \mathbb{1}_A((x_i, y_i)), \quad \forall A \in \mathcal{B}(\mathbb{R}^2).$$

Then we have $\text{LIS}(\xi) = \text{LIS}(\{(x_i, y_i)\}_{i=1}^n)$, where the latter one is defined in Definition 2.1.

Lemma 5.11. Let (V_1, \dots, V_n) be a random vector which has the same distribution as $(Y_1^{(n)}, \dots, Y_n^{(n)})$. For any $m \leq n$ and $\mathbf{b} = (b_1, b_2, \dots, b_m) \in Q(n, m)$, define the point process $\xi_{\mathbf{b}}^{(n)}$ as in the previous two lemmas, that is,

$$\xi_{\mathbf{b}}^{(n)}(A) := \sum_{i=1}^m \mathbb{1}_A((i, V_{b_i})), \quad \forall A \in \mathcal{B}(\mathbb{N} \times [0, 1]).$$

Then $\text{LIS}(\xi_{\mathbf{b}}^{(n)})$ and $\text{LIS}(\pi_{\mathbf{b}})$ have the same distribution, where $\pi \sim \mu_{n,q_n}$.

Proof. By the remarks after Definition 5.6, we have

$$\Phi(\{(i, V_{b_i})\}_{i=1}^m) = \Phi((V_{b_1}, V_{b_2}, \dots, V_{b_m})) = \Phi((V_1, V_2, \dots, V_n))_{\mathbf{b}}$$

where $\Phi((V_1, V_2, \dots, V_n))$ in the last term has the distribution μ_{n,q_n} . The lemma follows by the fact that

$$\text{LIS}(\xi_{\mathbf{b}}^{(n)}) = \text{LIS}(\{(i, V_{b_i})\}_{i=1}^m) = \text{LIS}(\Phi(\{(i, V_{b_i})\}_{i=1}^m)).$$

□

Now we are in the position to prove Lemma 4.1 and Lemma 4.2. In the following, we use λ_n to denote the uniform measure on S_n .

Proof of Lemma 4.1. The lemma can be divided into two parts. For the first part, we show that, for any sequence $\{k_n\}_{n=1}^\infty$ such that $k_n \in [n]$ and $\lim_{n \rightarrow \infty} k_n = \infty$,

$$\lim_{n \rightarrow \infty} \max_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} > 2 + \epsilon \right) = 0, \quad (82)$$

for any $\epsilon > 0$.

Since we have $q_n \geq 1$, by Corollary 3.5 (a), for any $\mathbf{b} \in Q(n, k_n)$, there exist two random variables (U, V) such that $U \sim \lambda_{k_n}$, V has the same distribution as $\pi_{\mathbf{b}}$ with $\pi \sim \mu_{n, q_n}$ and $U \leq_L V$. Hence we have $\text{LIS}(U) \geq \text{LIS}(V)$, since $\text{LIS}()$ is non-increasing on the poset (S_n, \leq_L) . Therefore, we have

$$\mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} > 2 + \epsilon \right) \leq \lambda_{k_n} \left(\pi \in S_{k_n} : \frac{\text{LIS}(\pi)}{\sqrt{k_n}} > 2 + \epsilon \right).$$

Then (82) follows by the result of Vershik and Kerov [14] that, under uniform measure, $\text{LIS}(\pi)/\sqrt{n}$ converges in probability to 2 as n goes to infinity. Note that (82) only depends on the fact that $q_n \geq 1$.

For the second part, we need to show that, for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \max_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} \leq 2e^{\frac{\beta}{2}} - \epsilon \right) = 0. \quad (83)$$

Given $n > 0$, for any $\mathbf{b} \in Q(n, k_n)$, by Lemma 5.11, $\text{LIS}(\xi_{\mathbf{b}}^{(n)})$ and $\text{LIS}(\pi_{\mathbf{b}})$ have the same distribution, where $\xi_{\mathbf{b}}^{(n)}$ is the point process as defined in that

lemma. Moreover, by Lemma 5.8, there exists a point process η_{k_n} such that $\eta_{k_n} \leq \xi_{\mathbf{b}}^{(n)}$ almost surely and η_{k_n} is defined by

$$\eta_{k_n}(A) := \sum_{i=1}^{k_n} B_{n,i} \cdot \mathbb{1}_A((i, U_i)) \quad \forall A \in \mathcal{B}(\mathbb{N} \times [0, 1]), \quad (84)$$

where $\{U_i\}_{i=1}^{k_n} \cup \{B_{n,i}\}_{i=1}^{k_n}$ are $2k_n$ independent random variables with each U_i being uniformly distributed on $[0, 1]$ and each $B_{n,i}$ being a Bernoulli random variable with $\mathbb{P}(B_{n,i} = 1) = 1/q_n^n$.

Hence, we have

$$\begin{aligned} \mu_{n,q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} \leq 2e^{\frac{\beta}{2}} - \epsilon \right) &= \mathbb{P} \left(\frac{\text{LIS}(\xi_{\mathbf{b}}^{(n)})}{\sqrt{k_n}} \leq 2e^{\frac{\beta}{2}} - \epsilon \right) \\ &\leq \mathbb{P} \left(\frac{\text{LIS}(\eta_{k_n})}{\sqrt{k_n}} \leq 2e^{\frac{\beta}{2}} - \epsilon \right). \end{aligned}$$

Here the last inequality follows by the monotonicity of $\text{LIS}()$ on \mathcal{X}_2 .

We complete the proof of (83) as well as Lemma 4.1 by showing the following,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\text{LIS}(\eta_{k_n})}{\sqrt{k_n}} > 2e^{\frac{\beta}{2}} - \epsilon \right) = 1, \quad (85)$$

for any $\epsilon > 0$.

From the inequality $\ln(1+x) \leq x$ for all $x > -1$, we have

$$\frac{1}{q_n^n} = e^{-n \ln q_n} \geq e^{-n(q_n-1)} = e^{n(1-q_n)}.$$

Since $\liminf_{n \rightarrow \infty} n(1-q_n) = \beta$, for any $\epsilon_1 > 0$, there exists $N_1 > 0$ such that, for any $n > N_1$, we have $1/q_n^n > e^{\beta-\epsilon_1}$. Thus, by the law of large numbers, we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} > k_n e^{\beta-\epsilon_1} \right) = 1. \quad (86)$$

Here we use the fact that $\lim_{n \rightarrow \infty} k_n = \infty$. Given $\mathbf{U} = (U_1, \dots, U_{k_n})$ and $\mathbf{B} = (B_{n,1}, \dots, B_{n,k_n})$, let $\Lambda(\mathbf{U}, \mathbf{B})$ denote the set of points in \mathbb{R}^2 defined by

$$\Lambda(\mathbf{U}, \mathbf{B}) := \{(i, U_i) : i \in [k_n] \text{ and } B_{n,i} = 1\}.$$

By the definition of η_{k_n} and Definition 5.10, we have

$$\text{LIS}(\eta_{k_n}) = \text{LIS}(\Lambda(\mathbf{U}, \mathbf{B})).$$

Moreover, conditioned on $\sum_{i=1}^{k_n} B_{n,i} = m$, by the independence of \mathbf{U} and \mathbf{B} , it is easily seen that $\text{LIS}(\Lambda(\mathbf{U}, \mathbf{B}))$ has the same distribution as $\text{LIS}(\pi)$ with $\pi \sim \lambda_m$.

For any $0 < \epsilon_2, \epsilon_3 < 1$, by the result of Vershik and Kerov [14] again, there exists $M > 0$ such that, for any $m > M$,

$$\lambda_m \left(\frac{\text{LIS}(\pi)}{\sqrt{m}} > 2 - \epsilon_2 \right) > 1 - \epsilon_3. \quad (87)$$

Since $\lim_{n \rightarrow \infty} k_n = \infty$ and (86), there exists $N > N_1$ such that, for any $n > N$, we have

$$k_n e^{\beta - \epsilon_1} > M \quad \text{and} \quad \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} > k_n e^{\beta - \epsilon_1} \right) > 1 - \epsilon_3.$$

Let $s = \lfloor k_n e^{\beta - \epsilon_1} \rfloor + 1$. For any $n > N$, we have

$$\begin{aligned} & \mathbb{P} \left(\text{LIS}(\eta_{k_n}) > (2 - \epsilon_2) \sqrt{k_n e^{\beta - \epsilon_1}} \right) \\ & \geq \sum_{m=s}^{k_n} \mathbb{P} \left(\text{LIS}(\eta_{k_n}) > (2 - \epsilon_2) \sqrt{k_n e^{\beta - \epsilon_1}} \mid \sum_{i=1}^{k_n} B_{n,i} = m \right) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & \geq \sum_{m=s}^{k_n} \mathbb{P} \left(\text{LIS}(\eta_{k_n}) > (2 - \epsilon_2) \sqrt{m} \mid \sum_{i=1}^{k_n} B_{n,i} = m \right) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & = \sum_{m=s}^{k_n} \lambda_m \left(\text{LIS}(\pi) > (2 - \epsilon_2) \sqrt{m} \right) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & > (1 - \epsilon_3) \sum_{m=s}^{k_n} \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & = (1 - \epsilon_3) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} > k_n e^{\beta - \epsilon_1} \right) \\ & > (1 - \epsilon_3)^2. \end{aligned}$$

Here the second inequality follows since $m \geq s \geq k_n e^{\beta - \epsilon_1}$. The third inequality follows from (87) and the fact that $m \geq k_n e^{\beta - \epsilon_1} > M$. Therefore, we have shown that $\lim_{n \rightarrow \infty} \mathbb{P} \left(\text{LIS}(\eta_{k_n}) > (2 - \epsilon_2) \sqrt{k_n e^{\beta - \epsilon_1}} \right) = 1$, and (85) follows from the fact that, by choosing ϵ_1 and ϵ_2 small enough, $(2 - \epsilon_2) \sqrt{e^{\beta - \epsilon_1}}$ can be arbitrarily close to $2e^{\frac{\beta}{2}}$. \square

The proof of Lemma 4.2 is similar to the proof of Lemma 4.1.

Proof of Lemma 4.2. Again, we split the proof into two parts. For the first part, we need to show that, for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \max_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} < 2 - \epsilon \right) = 0. \quad (88)$$

For this part, since $0 < q_n \leq 1$, we use Corollary 3.5 (b). For any $\mathbf{b} \in Q(n, k_n)$, there exist two random variables (U, V) such that $U \sim \lambda_{k_n}$, V has the same distribution as $\pi_{\mathbf{b}}$ with $\pi \sim \mu_{n, q_n}$ and $V \leq_L U$. Hence we have $\text{LIS}(V) \geq \text{LIS}(U)$, since $\text{LIS}()$ is non-increasing on the poset (S_n, \leq_L) . Therefore, we have

$$\mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} < 2 - \epsilon \right) \leq \lambda_{k_n} \left(\pi \in S_{k_n} : \frac{\text{LIS}(\pi)}{\sqrt{k_n}} < 2 - \epsilon \right).$$

Then (88) follows by the result of Vershik and Kerov [14] that, under uniform measure, $\text{LIS}(\pi)/\sqrt{n}$ converges in probability to 2 as n goes to infinity. Note that (88) only depends on the fact that $0 < q_n \leq 1$.

For the second part, we need to show that, for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \max_{\mathbf{b} \in Q(n, k_n)} \mu_{n, q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} \geq 2e^{\frac{\beta}{2}} + \epsilon \right) = 0. \quad (89)$$

First, we point out that, for any sequence $\{q_n\}_{n=1}^{\infty}$ with $0 < q_n \leq 1$ and $\limsup_{n \rightarrow \infty} n(1 - q_n) = \beta < \ln 2$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{q_n^n} = e^{\limsup_{n \rightarrow \infty} -n \ln q_n} = e^{\limsup_{n \rightarrow \infty} n(1 - q_n)} = e^{\beta} < 2.$$

Here the second equality follows from the fact that $\lim_{x \rightarrow 1} \frac{\ln x}{x-1} = 1$. Thus, for any $0 < \epsilon_1 < \ln 2 - \beta$, there exists $N_1 > 0$ such that, for all $n > N_1$, we have $1/q_n^n < e^{\beta + \epsilon_1}$.

Given $n > N_1$, for any $\mathbf{b} \in Q(n, k_n)$, by Lemma 5.11, $\text{LIS}(\xi_{\mathbf{b}}^{(n)})$ and $\text{LIS}(\pi_{\mathbf{b}})$ have the same distribution, where $\xi_{\mathbf{b}}^{(n)}$ is the point process as defined in that lemma. Moreover, by Lemma 5.9, there exists a point process ζ_{k_n} such that $\xi_{\mathbf{b}}^{(n)} \leq \zeta_{k_n}$ almost surely and ζ_{k_n} is defined by

$$\zeta_{k_n}(A) := \sum_{i=1}^{k_n} \mathbb{1}_A((i, U'_i)) + B_{n,i} \cdot \mathbb{1}_A((i, U_i)) \quad \forall A \in \mathcal{B}(\mathbb{N} \times [0, 1]), \quad (90)$$

where $\{U_i\}_{i=1}^{k_n} \cup \{U'_i\}_{i=1}^{k_n} \cup \{B_{n,i}\}_{i=1}^{k_n}$ are $3k_n$ independent random variables with each U_i, U'_i being uniformly distributed on $[0, 1]$ and each $B_{n,i}$ being a Bernoulli random variable with $\mathbb{P}(B_{n,i} = 1) = \frac{1}{q_n^n} - 1$.

Hence, we have

$$\begin{aligned} \mu_{n,q_n} \left(\pi \in S_n : \frac{\text{LIS}(\pi_{\mathbf{b}})}{\sqrt{k_n}} \geq 2e^{\frac{\beta}{2}} + \epsilon \right) &= \mathbb{P} \left(\frac{\text{LIS}(\xi_{\mathbf{b}}^{(n)})}{\sqrt{k_n}} \geq 2e^{\frac{\beta}{2}} + \epsilon \right) \\ &\leq \mathbb{P} \left(\frac{\text{LIS}(\zeta_{k_n})}{\sqrt{k_n}} \geq 2e^{\frac{\beta}{2}} + \epsilon \right). \end{aligned}$$

Here the last inequality follows by the monotonicity of $\text{LIS}()$ on \mathcal{X}_2 .

We complete the proof of (89) as well as Lemma 4.2 by showing that, for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\text{LIS}(\zeta_{k_n})}{\sqrt{k_n}} < 2e^{\frac{\beta}{2}} + \epsilon \right) = 1. \quad (91)$$

First, since, for all $n > N_1$, we have $\mathbb{P}(B_{n,i} = 1) = 1/q_n^n - 1 < e^{\beta+\epsilon_1} - 1$, by the law of large numbers, we get

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} < k_n(e^{\beta+\epsilon_1} - 1) \right) = 1. \quad (92)$$

Here we use the fact that $\lim_{n \rightarrow \infty} k_n = \infty$. Given $\mathbf{U}' = (U'_1, \dots, U'_{k_n})$, $\mathbf{U} = (U_1, \dots, U_{k_n})$ and $\mathbf{B} = (B_{n,1}, \dots, B_{n,k_n})$, let $\Lambda(\mathbf{U}', \mathbf{U}, \mathbf{B})$ denote the set of points in \mathbb{R}^2 defined by

$$\Lambda(\mathbf{U}', \mathbf{U}, \mathbf{B}) := \{(i, U_i) : i \in [k_n] \text{ and } B_{n,i} = 1\} \bigcup \{(i, U'_i) : i \in [k_n]\}.$$

By the definition of ζ_{k_n} and Definition 5.10, we have

$$\text{LIS}(\zeta_{k_n}) = \text{LIS}(\Lambda(\mathbf{U}', \mathbf{U}, \mathbf{B})). \quad (93)$$

Based on \mathbf{U}' , \mathbf{U} and \mathbf{B} , define another set of points in \mathbb{R}^2 as follows,

$$\Lambda^+(\mathbf{U}', \mathbf{U}, \mathbf{B}) := \{(i + 1/2, U_i) : i \in [k_n] \text{ and } B_{n,i} = 1\} \bigcup \{(i, U'_i) : i \in [k_n]\}.$$

Then, we have

$$\text{LIS}(\Lambda(\mathbf{U}', \mathbf{U}, \mathbf{B})) \leq \text{LIS}(\Lambda^+(\mathbf{U}', \mathbf{U}, \mathbf{B})). \quad (94)$$

Since, by Definition 2.1, no two points with the same x coordinates can be both within an increasing subsequence, by increasing the x coordinates of

those points in $\Lambda(\mathbf{U}', \mathbf{U}, \mathbf{B})$ which reside on the same vertical line as other points by $1/2$, the relative ordering of the shifted point with other points does not change, except the one which has the same x coordinate when unshifted. Combining (93) and (94), we have

$$\text{LIS}(\zeta_{k_n}) \leq \text{LIS}(\Lambda^+(\mathbf{U}', \mathbf{U}, \mathbf{B})). \quad (95)$$

Moreover, conditioned on $\sum_{i=1}^{k_n} B_{n,i} = m$, by independence of \mathbf{U}', \mathbf{U} and \mathbf{B} , it is easily seen that $\text{LIS}(\Lambda^+(\mathbf{U}', \mathbf{U}, \mathbf{B}))$ has the same distribution as $\text{LIS}(\pi)$ with $\pi \sim \lambda_{k_n+m}$. For any $0 < \epsilon_2, \epsilon_3 < 1$, by the result of Vershik and Kerov [14] again, there exists $M > 0$ such that, for any $k > M$,

$$\lambda_k \left(\frac{\text{LIS}(\pi)}{\sqrt{k}} < 2 + \epsilon_2 \right) > 1 - \epsilon_3.$$

Since $\lim_{n \rightarrow \infty} k_n = \infty$ and (92), there exists $N > N_1$ such that, for any $n > N$, we have

$$k_n > M \quad \text{and} \quad \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} < k_n(e^{\beta+\epsilon_1} - 1) \right) > 1 - \epsilon_3.$$

Let $s = \lceil k_n(e^{\beta+\epsilon_1} - 1) \rceil - 1$. For any $n > N$, we have

$$\begin{aligned} & \mathbb{P} \left(\text{LIS}(\zeta_{k_n}) < (2 + \epsilon_2) \sqrt{k_n e^{\beta+\epsilon_1}} \right) \\ & \geq \sum_{m=0}^s \mathbb{P} \left(\text{LIS}(\zeta_{k_n}) < (2 + \epsilon_2) \sqrt{k_n e^{\beta+\epsilon_1}} \mid \sum_{i=1}^{k_n} B_{n,i} = m \right) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & \geq \sum_{m=0}^s \mathbb{P} \left(\text{LIS}(\zeta_{k_n}) < (2 + \epsilon_2) \sqrt{k_n + m} \mid \sum_{i=1}^{k_n} B_{n,i} = m \right) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & \geq \sum_{m=0}^s \mathbb{P} \left(\text{LIS}(\Lambda^+(\mathbf{U}', \mathbf{U}, \mathbf{B})) < (2 + \epsilon_2) \sqrt{k_n + m} \mid \sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & \quad \times \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & = \sum_{m=0}^s \lambda_{k_n+m} \left(\text{LIS}(\pi) < (2 + \epsilon_2) \sqrt{k_n + m} \right) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & > (1 - \epsilon_3) \sum_{m=0}^s \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} = m \right) \\ & = (1 - \epsilon_3) \mathbb{P} \left(\sum_{i=1}^{k_n} B_{n,i} < k_n(e^{\beta+\epsilon_1} - 1) \right) \end{aligned}$$

$$> (1 - \epsilon_3)^2.$$

The second inequality follows because

$$k_n + m \leq k_n + s < k_n + k_n(e^{\beta+\epsilon_1} - 1) = k_n e^{\beta+\epsilon_1},$$

and the third inequality follows from (95). Therefore, we have shown that $\lim_{n \rightarrow \infty} \mathbb{P}\left(\text{LIS}(\zeta_{k_n}) < (2 + \epsilon_2)\sqrt{k_n e^{\beta+\epsilon_1}}\right) = 1$, and (91) follows from the fact that, by choosing ϵ_1 and ϵ_2 small enough, $(2 + \epsilon_2)\sqrt{e^{\beta+\epsilon_1}}$ can be arbitrarily close to $2e^{\frac{\beta}{2}}$. □

6 Discussion and open questions

A number of questions remain for further research.

1. In the proof of Corollary 1.5, we show that \bar{J} defined in Theorem 1.4 is attained when $\phi(x) = x$ given that $\lim_{n \rightarrow \infty} n(1 - q_n) = \lim_{n \rightarrow \infty} n(1 - q'_n) = \beta$. In fact, for any $\beta \in \mathbb{R}$, if $\gamma = 0, \beta, \pm\infty$, taking $\phi(x)$ to be the diagonal of the unit square gives the supremum of the following variational problem

$$\sup_{\phi \in B_{\gamma}^1} \int_0^1 \sqrt{\dot{\phi}(x) \rho(x, \phi(x))} dx.$$

Here, when $\gamma = \pm\infty$, we extend the definition of $\rho(x, y, \beta, \gamma)$ as follows, (We explicitly add β, γ as the argument of the density ρ .)

$$\rho(x, y, \beta, \pm\infty) := \lim_{\gamma \rightarrow \pm\infty} \rho(x, y, \beta, \gamma) = \lim_{\gamma \rightarrow \pm\infty} \int_0^1 u(x, t, \beta) \cdot u(t, y, \gamma) dt.$$

In fact, it is not hard to show that the above limits exist with

$$\rho(x, y, \beta, \infty) = u(x, y, \beta) \quad \text{and} \quad \rho(x, y, \beta, -\infty) = u(x, y, -\beta).$$

The question is whether $\phi(x) = x$ still solves the above variational problem for arbitrary $\beta, \gamma \in \mathbb{R}$.

2. The order of the expectation of the length of the longest common subsequence of two Mallows permutations under the same conditions as in Theorem 1.4 is still unknown to us. The main difficulty here is to bound the probabilities of those pairs of permutations (π, τ) such that $\text{LCS}(\pi, \tau) \gg 2\bar{J}\sqrt{n}$.
3. We are currently investigating the limiting behavior of the LCS between two Mallows permutations for the case when $\lim_{n \rightarrow \infty} n(1 - q_n) = \beta$ and $\lim_{n \rightarrow \infty} n(1 - q'_n) = \infty$ as well as the case when both limits are infinity. For the first case, it turns out that the weak law of large numbers of the LCS has exactly the same order as well as the leading constant as in the weak law of large numbers of the LIS of a Mallows permutation when $\lim_{n \rightarrow \infty} n(1 - q_n) = \beta$ (cf.[19]). The proof of this result is based on a coupling argument in which we exploit the stochastic dominance of Mallows measure under the weak Bruhat order. For the latter case in which both limits are infinity, we will adapt the techniques developed in [3] to show an analogous result under the following additional conditions,

$$\lim_{n \rightarrow \infty} q_n = \lim_{n \rightarrow \infty} q'_n = 1 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{1 - q_n}{1 - q'_n} = 1.$$

We will summarize these results in a future paper.

Acknowledgement

I am grateful to my supervisor Nayantara Bhatnagar for her helpful advice and suggestions during the research as well as her guidance in the completion of this paper.

References

- [1] James Abello, *The weak bruhat order of s_σ , consistent sets, and catalan numbers*, SIAM Journal on Discrete Mathematics **4** (1991), no. 1, 1–16.
- [2] Jinho Baik, Percy Deift, and Kurt Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, Journal of the American Mathematical Society **12** (1999), no. 4, 1119–1178.

- [3] Nayantara Bhatnagar and Ron Peled, *Lengths of monotone subsequences in a mallows permutation*, Probability Theory and Related Fields **161** (2015), no. 3-4, 719–780.
- [4] Václav Chvatal and David Sankoff, *Longest common subsequences of two random sequences*, Journal of Applied Probability (1975), 306–315.
- [5] Douglas E Critchlow, *Metric methods for analyzing partially ranked data*, vol. 34, Springer Science & Business Media, 2012.
- [6] Vladimír Dancík, *Expected length of longest common subsequences*, (1994).
- [7] Vlado Dančík and Mike Paterson, *Upper bounds for the expected length of a longest common subsequence of two binary sequences*, Random Structures & Algorithms **6** (1995), no. 4, 449–458.
- [8] Joseph G Deken, *Some limit results for longest common subsequences*, Discrete Mathematics **26** (1979), no. 1, 17–31.
- [9] Jean-Dominique Deuschel and Ofer Zeitouni, *Limiting curves for iid records*, The Annals of Probability (1995), 852–878.
- [10] Michael A Fligner and Joseph S Verducci, *Probability models and statistical analyses for ranking data*, vol. 80, Springer, 1993.
- [11] JM Hammersley, *A few seedlings of research in proc. of the sixth berkeley symp. math. statist. and probability, vol. 1 (1972), 345–394*, 1972.
- [12] Christian Houdré and Ümit Işlak, *A central limit theorem for the length of the longest common subsequence in random words*, arXiv preprint arXiv:1408.1559 (2014).
- [13] Ke Jin, *The limit of the empirical measure of the product of two independent mallows permutations*, arXiv preprint (2017).
- [14] SV Kerov and AM Vershik, *Asymptotics of the plancherel measure of the symmetric group and the limiting form of young tableaux*, Soviet Math. Dokl, vol. 18, 1977, pp. 527–531.
- [15] Benjamin F Logan and Larry A Shepp, *A variational problem for random young tableaux*, Advances in mathematics **26** (1977), no. 2, 206–222.

- [16] George S Lueker, *Improved bounds on the average length of longest common subsequences*, Journal of the ACM (JACM) **56** (2009), no. 3, 17.
- [17] Colin L Mallows, *Non-null ranking models. i*, Biometrika **44** (1957), no. 1/2, 114–130.
- [18] John I Marden, *Analyzing and modeling rank data, volume 64 of monographs on statistics and applied probability*, 1995.
- [19] Carl Mueller and Shannon Starr, *The length of the longest increasing subsequence of a random mallows permutation*, Journal of Theoretical Probability **26** (2013), no. 2, 514–540.
- [20] Pavel Pevzner, *Computational molecular biology: an algorithmic approach*, MIT press, 2000.
- [21] Shannon Starr, *Thermodynamic limit for the mallows model on s_n* , arXiv preprint arXiv:0904.0696 (2009).
- [22] Volker Strassen, *The existence of probability measures with given marginals*, The Annals of Mathematical Statistics (1965), 423–439.